

# 1 Large scale serum proteomics identifies proteins associated with 2 performance decline and clinical milestones in Duchenne muscular dystrophy

3 Authors:

4 N.A. Ikelaar<sup>1,2,\*</sup>, A.M. Barnard<sup>3,\*</sup>, S.W.M. Eng<sup>4</sup>, S. Hosseini Vajargah<sup>4</sup>, K.C.H. Ha<sup>4</sup>, H.E. Kan<sup>2,5</sup>, K.  
5 Vandenborne<sup>3</sup>, E.H. Niks<sup>1,2</sup>, G.A. Walter<sup>6,§</sup>, P. Spitali<sup>7,§</sup>

## 7 Affiliations:

- 8 1. Department of Neurology, Leiden University Medical Center, Leiden, Zuid-Holland, the
- 9 Netherlands
- 10 2. Duchenne Center Netherlands
- 11 3. Department of Physical Therapy, University of Florida, Gainesville, FL, USA
- 12 4. BioSymetrics, Inc., Huntington, NY, USA.
- 13 5. C.J. Gorter MRI Center, Department of Radiology, Leiden University Medical Center, Leiden, Zuid-
- 14 Holland, the Netherlands
- 15 6. Department of Physiology and Aging, University of Florida, Gainesville, FL, USA
- 16 7. Human Genetics Department, Leiden University Medical Center, Leiden, Zuid-Holland, the
- 17 Netherlands

18  
19 \* Authors contributed equally

20 § Authors contributed equally

21

22

23

24

25

26

27

28

29

30

31

32

33

34

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

35

36

37 **Abstract**

38 Serum biomarkers are promising minimally invasive outcome measures in clinical studies in  
39 Duchenne muscular dystrophy (DMD). However, biomarkers strongly associated with clinical  
40 progression and predicting performance decline are lacking. In this study we aimed to identify serum  
41 biomarkers associated with clinical performance and able to predict clinical milestones in DMD.  
42 Towards this aim we present a retrospective multi-center cohort study including serum samples and  
43 clinical data collected in research participants with DMD as part of a natural history study at the  
44 University of Florida (UF) and real-world observations at Leiden University Medical Center (LUMC)  
45 between 2009-2022. The 7K SomaScan® assay was used to analyse protein levels in individual  
46 serum samples. Serum biomarkers predicted age at loss of ambulation (LoA), age at loss of overhead  
47 reach (OHR) and age at loss of hand to mouth function (HTM). Secondary outcomes were the  
48 association of biomarkers with age, corticosteroid (CS) usage, and clinical performance based on the  
49 North Star Ambulatory Assessment (NSAA), 10 meter run velocity (10mrv), 6 minute walk (6MWT)  
50 and Performance of the Upper Limb (PUL2.0). A total of 716 serum samples were collected in 79  
51 participants at UF and 74 at LUMC (mean[SD] age; 10.9[3.2] vs 8.4[3.4]). 244 serum proteins showed  
52 an association with CS usage in both cohorts independent of CS type and regimen, including MMP3  
53 and IGLL1. 318 probes (corresponding to 294 proteins) showed significant associations with NSAA,  
54 10mrv, 6MWT and/or PUL2.0 across both cohorts. The expression of 38 probes corresponding to 36  
55 proteins such as RGMA, EHMT2, ART3, ANTXR2 and DLK1 was associated with risk of both lower and  
56 upper limb clinical milestones in both the LUMC and UF cohort. In conclusion, multiple biomarkers  
57 were associated with CS use, motor function and upper lower and upper limb disease milestones in  
58 DMD. These biomarkers were validated across two independent cohorts, increasing their likelihood  
59 of translation for use within the broader DMD population.

60

61

62

63

64

65

66

67

68

69

70

## 71 Introduction

72 Duchenne muscular dystrophy (DMD) is the most common form of muscular dystrophy<sup>1</sup>. It is caused  
73 by variants in the *DMD* gene, resulting in an absence of functional dystrophin<sup>2</sup>. Symptoms begin in  
74 early childhood with delayed motor milestones and continue with progressive muscle weakness,  
75 scoliosis, reduced pulmonary function, and dilated cardiomyopathy. Untreated, most patients will  
76 lose ambulation by the age of 10 years<sup>3,4</sup>.

77 Current treatment consists of long-term corticosteroid (CS), initiated at age 4-5 years and  
78 recommended for life-long use<sup>5</sup>. Chronic use of CS has been shown to delay loss of ambulation,  
79 onset and progression of scoliosis, and, in combination with non-invasive ventilation, it increases life  
80 expectancy to 30-40 years of age<sup>3,4</sup>. Yet CS are not without negative metabolic effects<sup>6</sup>. Despite the  
81 huge progress in understanding the biology of the disease and the large investment in pre-clinical  
82 research, there is currently no cure for DMD. Multiple investigational drugs have failed to show  
83 significant improvement in participants' performance in clinical trials, with only a few drugs receiving  
84 full and/or accelerated or conditional FDA approval such as microdystrophin gene therapy<sup>7</sup>, exon  
85 skipping<sup>8</sup>, vamorolone<sup>9</sup> and givinostat<sup>10</sup>. While failures may be related to limited drug potency, it has  
86 also become clear that disease trajectories present high inter-individual variation, further reducing  
87 the power to detect a significant and clinically meaningful treatment effect on a single outcome  
88 measure in a typically short term, multi-center clinical trial. Several factors contribute to disease  
89 heterogeneity, including mutation type<sup>11-14</sup>, genetic modifiers<sup>15</sup>, and standards of care. Outcome  
90 measures have been shown to be affected by patient motivation<sup>16</sup>. These factors complicate  
91 accurate prediction of individual disease trajectories and thereby contribute to difficulty in clinical  
92 trial design. Therefore, being able to objectively predict long-term clinical outcomes based on short  
93 term evidence would greatly facilitate the evaluation of medicinal products.

94 Numerous serum biomarkers have been proposed for DMD<sup>17-19</sup>, however the focus has mostly been  
95 in the identification of cross-sectional differences between unaffected controls and individuals with  
96 DMD. Monitoring biomarkers, evaluated in series, are particularly lacking. Research efforts focused  
97 on the analysis of muscle damage biomarkers<sup>20-22</sup>, and in monitoring response to (micro-)dystrophin  
98 restoration therapies<sup>23,24</sup>. However, serum biomarkers with prognostic value are currently lacking.  
99 The challenge and paradox relies on a disconnect between the early window of opportunity to treat  
100 patients (roughly 3-8 years of age, in which most of the muscle mass is available for therapeutic  
101 applications aimed at prevention vs reversal), the observable decline in patients function (typically  
102 starting around 8-9 years of age), and the observation of disease milestones such as loss of  
103 ambulation (between 10-16 years of age). Initial attempts to determine longitudinal trajectories in  
104 serum biomarkers related to disease progression<sup>22,25,26</sup> have been limited by the small number of  
105 individuals, lack of follow-up samples, and incomplete collection of clinical data.

106 Therefore, discovery studies in large and well-characterized cohorts are needed to identify  
107 associations between biomarker signatures and clinical performance and to define the context of  
108 use for the biomarker signature. Importantly, prognostic biomarkers that can predict disease  
109 milestones over longer time frames could enrich clinical trial design. Patients recruitment can be  
110 tailored to individuals who are more likely to experience the milestone during the study. This is  
111 particularly relevant as participants are usually enrolled in clinical trials before the age of 8, in early  
112 stages of disease progression, but when little decline can be normally observed within the duration

113 of an interventional study of 12 months. Connecting short term changes in biomarker signatures to  
114 the likelihood of long-term decline would help bridge this gap. Therefore, in this study we aimed to  
115 identify serum biomarkers associated with clinical outcomes and their prognostic value for clinically  
116 relevant milestones for both lower and upper limb function.

117

118

## 119 **Methods**

### 120 Study Cohort, design and outcomes

121 This was a retrospective, multi-center, cohort study including serum samples and clinical data  
122 collected from individuals with DMD participating in research protocols at the University of Florida  
123 (UF) and at the Leiden University Medical Center (LUMC) between 2009-2022. We included 407  
124 serum samples from 74 individuals with DMD aged 4 – 24 years at LUMC and 295 serum samples  
125 from 79 individuals with DMD aged 5 – 22 years at UF. At LUMC, samples were collected during  
126 yearly visits to the outpatient clinics as part of the standard of care. At UF, samples were collected as  
127 part of an optional biosample addition to research visits conducted for the ImagingDMD natural  
128 history study (NCT01484678). Written consent was obtained from all participants or their caregivers  
129 as described in protocol B22.013 at LUMC and protocols IRB201500981 and IRB201700056 at UF,  
130 which were approved by the respective regulatory boards at both sites.

131 Clinical data were obtained at the same clinic or research visit as serum sample collection. Data  
132 included age at sample collection, CS use at the time of sample collection, and performance on tests  
133 of function. CS information was categorized by use (treated or untreated), type (deflazacort,  
134 prednisone, or other), regimen (daily or intermittent – defined as 10 days on/10 days off or weekend  
135 dosing), and dose. Motor function tests included the North Star Ambulatory Assessment (NSAA), 10  
136 meter run/walk velocity (10MRV), 6 minute walk test (6MWT), and Performance of the Upper Limb  
137 (PUL2.0)<sup>27–30</sup>. Three disease milestones were recorded: age at loss of ambulation (LoA), age at loss of  
138 overhead reach (OHR), and age at loss of hand to mouth (HTM). LoA was defined similarly in both  
139 cohorts with LoA based on patient reported inability to walk 5 meters unaided at home at LUMC and  
140 inability to traverse 10 meters unaided within 45 seconds at UF. OHR and HTM were primarily  
141 derived from PUL2.0 scores and occasionally from patient reported data.

142

### 143 Samples collection and proteomic analysis

144 Serum samples were collected and prepared according to standard phlebotomy procedures.  
145 Samples at LUMC were left to clot for ~30 minutes, followed by 10 min centrifugation at 2350g,  
146 while clotting time was 30 minutes at UF, followed by centrifugation at 3000 rpm for 15 min. At the  
147 LUMC, sample aliquots were frozen at -20°C for 1-2- months and then transferred to -80°C for long  
148 term storage, while at UF, samples were immediately stored at -80°C. Serum samples were analysed  
149 by the SomaScan® proteomic platform at SomaLogic (Boulder, Colorado, USA)<sup>31</sup>. SomaScan uses  
150 aptamers as affinity reagents to detect and quantify proteins in complex mixtures such as serum.  
151 Aptamers are nucleic acids that are able to bind specific protein targets with high specificity and  
152 affinity. The pipeline is built to keep targets engaging aptamers, while washing away the unbound  
153 ones. Detection of the aptamers is then performed using an array with probes complementary to the  
154 aptamer sequence. At the time of this research, this platform included 7596 aptamers capable of  
155 detecting 6628 proteins. As part of the assay, extensive quality control metrics are put in place  
156 including internal and external controls. A total of 9 samples (2 LUMC / 7 UF) did not pass  
157 SomaLogic's internal quality control and were excluded from further analysis, results, and data  
158 tables.

159 Statistical analysis

160 Linear mixed effects models (LMEM) were constructed for each site and probe to predict that  
161 probe's expression given age and CS use as follows:

162 standardized log-transformed probe expression ~ standardized age, grouped by patient

163 standardized log-transformed probe expression ~ standardized age + corticosteroid use (yes/no),  
164 grouped by patient

165 LMEMs were then used to predict motor function tests, controlling for age and CS use as follows:

166 standardized motor function test performance ~ standardized log-transformed probe expression +  
167 standardized age + corticosteroid use (yes/no), grouped by patient

168 LMEMs were conducted using the *pymr4* package, version 0.7.8, a Python interface for the *lme4* R  
169 package, version 1.1-31<sup>32,33</sup>. To determine how generalizable the significant associations between  
170 probes and motor function tests were, for each site, the associated LMEMs were used to predict  
171 values of the motor function tests for the other site. That is, for LUMC models, UF predicted values  
172 were generated using UF data, and for UF models, LUMC predicted values were generated using  
173 LUMC data. Predicted values were compared against original values via reconstruction accuracy:

174 
$$Q^2 = 1 - \frac{\sum(\text{predicted values} - \text{original values})^2}{\sum(\text{predicted values})^2}$$

175

176 Significant probes whose validation  $Q^2$  values were within the top 5% of  $Q^2$  values (i.e., at or above  
177 the 95<sup>th</sup> percentile) were kept for prediction of the clinical milestones (age at LoA, loss of OHR, and  
178 loss of HTM). These probes were grouped based on the milestone being predicted for each site. For  
179 LoA, probes associated with NSAA, 10MRV, and/or 6MWT were considered. For loss of HTM and  
180 OHR, probes associated with PUL2.0 were considered. For each site, clinical milestone, and  
181 candidate probe, a Cox proportional hazards model was constructed to predict that milestone using  
182 the following specification:

183 clinical milestone ~ standardized log-transformed probe expression + corticosteroid use (yes/no),  
184 with age as the time component and grouped by patient.

185 Modelling was conducted using the *lifelines* Python package, version 0.27<sup>34</sup>. All p-values underwent  
186 false discovery rates (FDRs) adjustments to account for multiple hypothesis testing. Probes with  
187 FDRs < 0.05 were considered to be significantly associated with the respective outcomes. A probe  
188 was significantly associated with an outcome if its FDR was <0.05.

189 Data availability

190 Anonymized data can be made available to qualified investigators on request. Requests should be in  
191 line with the approved ethical protocol.

192

193 **Results**

194 After retaining samples passing quality control, a total of 153 males with DMD and 693 serum  
 195 samples were included in the study (Table 1). LUMC participants were significantly younger at the  
 196 first sample visit compared to UF participants (mean[SD]; 8.4[3.4] vs 10.9[3.2],  $p < 0.001$ ). An average  
 197 of 4.3 samples per participant were analysed, with longer follow-up duration in the LUMC cohort  
 198 (5.7[3.6] vs 3.4[2.5] years,  $p < 0.001$ ). An intermittent (10 days on/10 days off of pred (or Deflazacort)  
 199 CS regimen was most common at LUMC while daily dosing was most common at UF (prednisone and  
 200 deflazacort); however, age at start of CS treatment was comparable across sites. Only a few patients  
 201 remained CS naïve for the entire study period (4.1% vs 3.2%). The percentage of patients meeting  
 202 one of the 3 clinical milestones at first sample was higher at LUMC compared to UF (LoA: 18.9% vs  
 203 9.6%, loss of OHR: 16.0% vs 12.2%, loss of HTM: 5.3% vs 0.0%). For the LUMC cohort, 6660 probes  
 204 passed quality control compared to 6690 for the UF cohort.

205

	LUMC n = 74	UF n = 79
<i>Samples, n</i>	405	288
<i>Age at first sample, years, mean (SD)</i>	8.4 (3.4)	10.9 (3.2)
<i>Follow-up duration, years, mean (SD)</i>	5.7 (3.6)	3.4 (2.5)
<i>Samples per patient, n, mean (SD)</i>	5.5 (2.9)	3.6 (2.0)
<i>Age at start of corticosteroids, years, mean (SD)</i>	5.6 (1.6)	5.4 (1.7)
<i>Corticosteroid treatment at first sample, n (%)</i>		
• Not recorded	1 (1.4%)	3 (3.8%)
• Daily	2 (2.7%)	66 (83.5%)
• Intermittent	49 (66.2%)	4 (5.1%)
• None	22 (29.7%)	6 (7.6%)
<i>Dystrophin variants, n (%)</i>		
• Exonic deletion	49 (66.2%)	49 (62.0%)
• Exonic duplication	10 (13.5%)	8 (10.1%)
• Nonsense variant	13 (17.6%)	12 (15.2%)
• Splice site variant	2 (2.7%)	3 (3.8%)
• Other small point variants	—	5 (6.3%)
• Variant unknown	—	2 (2.5%)
<i>Weight at first sample, kg, mean (SD)</i>	32.6 (18.6)	35.1 (12.2)
<i>Height at first sample, cm, mean (SD)</i>	127.2 (19.2)	126.9 (10.7)
<i>BMI at first sample, mean (SD)</i>	18.3 (4.4)	21.2 (5.4)
<i>Ambulant patients at first sample, n (%)</i>	59 (79.7%)	68/78 (87.2%) (1 unknown)
<i>Able to perform overhead reach at first sample, n (%)</i>	64 (86.5%)	63/70 (90.0%) (9 unknown)
<i>Able to perform hand to mouth at first sample, n (%)</i>	73 (98.6%)	74/75 (98.7%) (4 unknown)

206 *Table 1. Cohort characteristics.*

207 Serum protein associations with age and corticosteroid usage

208 Given the progressive nature of DMD and the availability of longitudinal visits, we first identified  
 209 proteins associated with age. 4,796 probes (4436 proteins) were significantly associated with age in  
 210 the LUMC cohort and 2,668 probes (2498 proteins) were significantly associated with age in the UF  
 211 cohort (FDR<0.05)(Fig. 1A-B); with overlap of 2,317 probes (2186 proteins) between both cohorts  
 212 (Fig. 1C). 2,251 probes showed concordant directional change between cohorts, with the  
 213 coefficients generally being higher in the LUMC cohort (Fig. 1D). Multiple muscle proteins were

214 negatively associated with age including, but not limited to, CK-MM, MYOM3, MYOM2, TRIM72,  
215 MYL3, BIN1, TTN, LAMA2, TPM3, ACTN2, DES, CA3, MYL3 and TNNI2 (Fig1E-H). Markers of  
216 fibroadipogenic progenitors (PDGFRA), fibrosis and extracellular matrix (collagens such as COL1A1,  
217 COL2A1, COL3A1, COL6A1, COL6A2, COL6A3, TIMP1), several BMPs (BMP-1,4,5,6,7) and protein  
218 synthesis pathways (AKT1, AKT2, IGF2), also decreased with increasing age.

219 Among the proteins with expression that significantly increased with age, we identified proteins  
220 present in the central nervous system such as CNTN3 and CNDP1, proteins expressed in adipose  
221 tissue (LEP, GHR, FABP4, ADIPOQ, CNTFR, CFD and TIMP4, PNLIPRP1, PNLIPRP2, PNLIP), and several  
222 cytokines (CCL14, IFNA6, BMPR2, CXCL8, CCL3L1, IL26, IL5RA, CCL1, CCL16, CCL15, IL10RA, IL31RA,  
223 IL17RE) (Fig1I-L). Two different probes for IGF1 had had opposing associations with age, and several  
224 IGF binding proteins showed discordant trajectories, with IGFBP1 and IGFBP2 decreasing with age  
225 and IGFBP5 and IGFBP6 increasing with age. Coefficients of all significant, concordant proteins  
226 associated with age are provided in Supplementary Table 1.

227 Given that the majority of participants had chronic exposure to CS , we sought to identify a protein  
228 expression signature related to exposure to CS while correcting for age. 846 probes (790 proteins) in  
229 the LUMC cohort and 396 probes (364 proteins) in the UF cohort were significantly associated with  
230 CS use (FDR < 0.05)(Fig. 2A-B). Furthermore, 244 probes (227 proteins) were shared across sites (Fig.  
231 2C). The number of proteins associated with steroid use in the UF cohort was smaller due to the  
232 small number of patients untreated (n=5) in this cohort compared to the untreated group in the  
233 LUMC cohort (n=22), suggesting that the number of proteins identified as significantly associated  
234 with the treatment is not a measure of the treatment effect. This is confirmed by the comparison of  
235 the coefficients shown in Fig.2D, showing larger effects for the UF cohort in line with the higher  
236 dosage (mostly daily) compared to the LUMC cohort (mostly intermittent). Among the proteins we  
237 identified (CD23) and (MMP3)<sup>6</sup> previously observed in DMD, as well as newly identified proteins  
238 such as IGLL1 immunoglobulin lambda-like polypeptide 1 (IGLL1) and repulsive guidance molecule A  
239 (RGMA)(Fig. 2E-G). To identify whether the identified associations were potentially related to  
240 efficacy or safety effects, we assessed the directionality of the signature for age and steroids.  
241 Proteins showing discordant association (such as declining with age as disease progresses and  
242 increased by CS treatment) were considered to relate to the steroid efficacy signature, while  
243 concordant effects were considered to belong to the safety signature (Fig.2I). A total of 44 proteins  
244 showed discordant age and CS effects (potential efficacy signature) across the 2 cohorts  
245 (Supplementary table 2); this list included previously identified proteins such as ANGPT2 (2  
246 independent aptamers), proteins related to lipoproteins transportation (APOA4, APOC3, APOE) as  
247 well as newly identified recognized by 2 independent aptamers such as DLD, ART3, NDUFA2, OLFM2  
248 as well as 2 members of the RGM family (RGMA and RGMC)(Fig.2H). A total of 172 proteins showed  
249 concordant sign for age and CS effects in both LUMC and UF cohorts (potential safety biomarkers;  
250 Supplementary Table 3). This list included previously identified MMP3, Afamin, IGFBP-5, proteins  
251 involved in SMAD signalling as well as new proteins belonging with IGF binding properties and  
252 metalloproteinases.

253

254 Serum protein associations with motor performance



255 Longitudinal trajectories for the NSAA, 10MRW, 6MWT, and PUL 2.0 differed between the two sites,  
256 with a shift towards earlier decline in motor performance in the LUMC cohort (Fig 3A-D). To assess  
257 which proteins were associated with clinical severity, we evaluated associations with each functional  
258 test in each cohort. 2,005 probes (1882 proteins) showed a significant association with at least one  
259 motor function test in the LUMC cohort compared to 483 probes (454 proteins) in the UF cohort  
260 (FDR < 0.05). 318 probes (294 proteins) were shared across the two sites, with the majority of  
261 significant associations found for lower limb scales, consistent with the higher number of  
262 observations and the larger degree of functional decline experienced in the cohorts. Cross-  
263 correlation of the coefficients showed concordant directionality between the 2 cohorts, especially  
264 for the lower limb scales (Fig. 3E).

265 We then assessed how generalizable models trained in each cohort were by using the other cohort  
266 for validation, selecting the 5% most performant proteins for each scale. Validation reconstruction  
267 accuracies were higher for lower limb function with  $Q^2$  values comparable across both cohorts for  
268 the 10MRV, while accuracies were higher for models trained on the LUMC cohort for the 6MWT and  
269 higher for UF for the NSAA. 122 probes (116 proteins) in the LUMC cohort and 109 probes (105  
270 proteins) in the UF cohort significantly and accurately predicted all 3 scales used to monitor lower  
271 limb function (Fig. 3F). Supplementary Table 4 provides a list of proteins with significant associations  
272 across scales and cohorts. ART3, RGMA and DLD showed highly significant associations and accurate  
273 predictions across multiple motor function tests (Fig. 3G-H). To refine the efficacy and safety  
274 signature identified in the steroid analysis we assessed whether cross-correlation between steroid  
275 and clinical association coefficients; we identified 919 pairs where a protein was significantly  
276 associated with a clinical score and with steroid treatment across both cohorts. Concordant  
277 coefficients are in this case expected to be considered efficacy biomarkers, while discordant  
278 coefficients are expected to be adverse effects (Fig. 3I). A chi-squared test showed that concordant  
279 signs were significantly enriched in the previously defined efficacy signature, while discordant signs  
280 were enriched in the safety signature ( $P < 10^{-15}$ ), with only 835 of the 919 pairs being properly  
281 assigned.

## 282 Prediction of clinical milestones

283 Given the association with continuous clinical scores, we sought to determine whether the identified  
284 proteins had prognostic value for clinically meaningful disease milestones such as age at LoA, loss of  
285 OHR, and loss of HTM. All clinical milestones were reached at significantly earlier ages and in a  
286 greater percentage of patients in the LUMC cohort than in the UF cohort (Table 1, Fig. 4A-B). We  
287 identified 41 probes / milestone associations. Among these, 3 of the probes targeting RGMA, ENPP5  
288 and RGS21 were associated with 2 different milestones (FDR < 0.05). RGMA was associated with  
289 lower and upper limb milestones across the 2 cohorts, , while RGS21 (UF) and ENPP5 (LUMC) were  
290 cohort-specific. Associations for DLK1 and ART3 were confirmed by 2 independent probes: DLK1 was  
291 found to be associated with OHR in the UF cohort and ART3 with LOA in the LUMC cohort (Fig.4C).  
292 The direction of log Hazard Ratios (logHR) were conserved for these proteins across probes and  
293 outcomes. Notably, RGMA, ANTXR2, EIF4G1, CAMK2A had large negative logHRs and PTPRD, NCAM  
294 had large positive logHRs. Kaplan-Meier plots for RGMA are presented across all 3 milestones for  
295 both LUMC and UF cohorts (Fig 4D). Supplementary Table 5 lists all proteins significantly predicting  
296 lower and upper limb milestones. Since treatment with steroids is known to delay these milestones,  
297 we assessed whether logHR and steroid effects show discordant effects. Figure 4E illustrates how

298 estimates are indeed anti-correlated for all but 2 proteins, supporting the use of these biomarkers to  
299 evaluate the risk of meeting these milestones.

300 Among the multiple proteins associated with functional tests and disease milestones, we shortlisted  
301 a few based on the consistency and strength of the associations across cohorts (Table 2). Proteins  
302 ANTXR2, ART3, EHMT2 and RGMA showed a % risk increase for disease milestones between 136%  
303 and 981% for a decrease of one unit (in log scale). Treatment with steroids corresponded to an  
304 increase in the biomarker level corresponding to a reduced the risk % up to 90% with highest  
305 normalization of risk for the upper limb HTM milestone; normalization for the risk % of LoA was in all  
306 cases estimated between 22% and 39%. We further estimated how monitoring proteins in blood can  
307 help evaluate the yearly increase in risk % for the milestones and observed yearly risk increase  
308 between 43% and 215%.

Gene Symbol	Probe	Site	Milestone	Log HR	$\beta$ CS	$\beta$ age	Risk % increase with 1 log decrease	Risk % reduction with CS treat	Yearly risk % increase
ANTXR2	15559-5	LUMC	LoA	-1.14	0.43	-0.76	213.54	38.50	138.38
ART3	7970-315	LUMC	LoA	-0.89	0.47	-0.41	144.55	34.49	43.83
ART3	10970-3	LUMC	LoA	-0.86	0.30	-0.48	135.79	22.67	51.19
EHMT2	5843-60	LUMC	LoA	-1.02	0.47	-0.60	176.06	38.11	83.08
RGMA	5483-1	UF	HTM	-2.38	0.98	-0.48	981.79	90.38	215.03
RGMA	5483-1	LUMC	LoA	-1.00	0.51	-0.64	170.79	39.78	89.96

309

310 *Table 2. Shortlisted proteins with consistent association across performance tests, milestones and*  
311 *cohorts. The coefficient estimates for the indicated milestone are obtained from the time to event*  
312 *analysis, while the coefficient for CS and age were obtained by the linear mixed effect models*  
313 *including both CS and age. We further show the risk % for the indicate milestone for a unit decrease*  
314 *in the protein (in log scale) and how the risk is reduced by treatment with CS and increased per year.*

315

316

## 317 Discussion

318 Large-scale serum biomarker discovery is now possible in DMD with the availability of high-  
319 throughput proteomics platforms. In this retrospective study, we had the unique opportunity to pair  
320 robust and comprehensive clinical data from two large, independent international cohorts of  
321 individuals with DMD to serum levels of 6,628 proteins. The availability of extensive longitudinal  
322 serum samples and clinical data not only allowed for analysis of protein signatures associated with  
323 age and corticosteroid usage, but also allowed for identification of novel proteins predictive of  
324 clinical function and clinical milestones. Proteins like RGMA were predictive of both lower and upper  
325 limb clinical milestones such as LoA and loss of HTM. Furthermore, several proteins were associated  
326 with increased risk of single clinical milestones such as ART3 for LoA. The ability to compare these  
327 results across the two independent cohorts enabled for cross-validation of our findings.

328 Prior serum biomarker discovery in dystrophinopathies has largely been limited to smaller cohorts,  
329 targeted protein selection for analysis, and limited longitudinal data. These studies have been crucial  
330 for building our understanding of the disease's pathogenic processes. Key differences between  
331 individuals with and without DMD have been reported, including several proteins associated with  
332 age, shedding light on the ongoing pathogenic processes. It is well-known with DMD that serum CK  
333 (activity and protein levels) decrease with age, reflecting muscle damage and loss of muscle mass.  
334 Additional proteins decreasing with age have subsequently been identified including MDH2, MYL3,  
335 CA3, MYOM3, COL1A1, TTN, TNNT2, ETFA, TNNT3, and MAP4<sup>26,35</sup>. Proteins that have increasing  
336 expression with age have included C4A, C4BPA, and GSN, among others<sup>35</sup>. A small number of studies  
337 have also utilized the SomaScan® platform to perform large-scale protein biomarker analysis in  
338 DMD, identifying additional proteins associated with age/time such as LEP, CFTNR, FABP3, and  
339 TNNT2<sup>25,36</sup>. Complementing and largely validating the findings of these studies, we identified  
340 previously described proteins such as CK and MYOM-3 showing steep decline in protein levels, but  
341 also substitution of muscle with fibro-adipogenic tissue with markers such as Collagen 1 and 2  
342 declining along with markers of fibro-adipogenic progenitors such as PDGFRA<sup>37</sup>. We also observed  
343 adipogenic markers such as LEP, GHR and ADIPOQ with increased expression along with a  
344 complement and inflammation signature.

345 The CS signature encompassed steroid engagement biomarkers such as MMP3 and IGLL1, which  
346 showed the strongest association with CS usage, both in the daily and the intermittent regimen.  
347 Previous research has also identified elevated MMP3 in patients with DMD treated with CS, while  
348 IGLL1 has not yet been described in relation to DMD<sup>38</sup>. However, it has very recently been shown  
349 that weekend CS use in adults with limb girdle muscular dystrophy reduces IGLL1 in the presence of  
350 MMP3 elevation<sup>39</sup>. Given the role of IGLL1 in B cells, the reduction observed in combination with the  
351 CS treatment could signal the immunosuppression expected in these patients due to the drug.

352 Treatment with corticosteroids contributed to normalize as well as exacerbate the disease  
353 progression signature. An exacerbation of the signature related to age was observed for previously  
354 reported safety biomarkers such as MMP3, Afamin, and IGFBP5<sup>5</sup>, and for certain apolipoproteins  
355 such as APOA2, APOL1, APOA5. A normalization was observed for other apolipoproteins such as  
356 APOE4, APOC3 and APOE, suggesting that steroid treatment could directly normalize the  
357 dyslipidemia with effects on APOE4, APOC3, and APOE, while exacerbation of the APOA2, APOL1,  
358 and APOA5 signature could highlight how steroids affect lipid metabolism and potentially

359 cardiovascular health. The compensatory effect of APOE is also underscored by the more severe  
360 phenotype observed in *mdx* ApoE double knock-out mice<sup>40,41</sup>. We also found that other previously  
361 reported proteins such as ANGPT2<sup>6</sup>, and proteins (RGMA, DLK1, ANTXR2, and ART3) that were  
362 decreased with disease progression in both UF and Leiden cohorts were higher in the CS groups  
363 compared to untreated.

364 A major strength of this study was the ability to identify serum biomarkers predictive of clinical  
365 function. RGMA and ART3 were all directly related to muscle performance as measured by both  
366 upper (PUL2.0) and lower limb (NSAA, 6MWT, 10MRV) functional outcomes. RGMA, ANTXR2,  
367 EHMT2 and ART3 were found to have large negative logHR when considering the loss of certain  
368 clinical milestones. To further illustrate this, we stratified the population according to RGMA levels  
369 and found that the decreasing levels of RGMA clearly shifted the Kaplan-Meier curve towards earlier  
370 LoA and OHR.

371 RGMA is part of the repulsive guidance molecule family of glycoprotein-1 (GP1) anchor proteins that  
372 is mostly expressed in the central nervous system and muscle tissue according to the human protein  
373 atlas gene expression data. Originally described as playing an important role in neurogenesis by  
374 guiding axonal growth and as an important target of neuronal survival, it has now also been  
375 identified to play a role in myogenesis<sup>42,43</sup>. RGMA has been proposed to play a central role in  
376 regulating cellular hypertrophy and hyperplasia<sup>44</sup>. RGMA has previously been found to have an  
377 association with Spinal and bulbar muscular atrophy (SBMA)<sup>45</sup>, Parkinson disease, Alzheimer disease,  
378 multiple sclerosis, and cerebrovascular accidents, as well as an association with upper limb function  
379 measured by elbow flexion in DMD<sup>25,42</sup>.

380 ANTXR2, also known as capillary morphogenesis protein 2 (CMG2) plays an important role in cellular  
381 interaction by binding collagen IV and laminin, suggesting involvement in extracellular matrix  
382 adhesion. It is expressed in multiple tissues, including muscles. Loss of function variants in ANTXR2  
383 cause hyaline fibromatosis syndrome and ANTXR2 knockout mice show collagen VI accumulation in  
384 the uterus<sup>46</sup>. This could implicate ANTXR2 as playing a role as a collagen VI regulator, and possible  
385 involvement in muscle homeostasis.

386 ART3, known as ADP-Ribosyltransferase 3, is mainly expressed in skeletal muscle tissue according to  
387 GTEx and human protein atlas databases. Gene expression data obtained from human individuals as  
388 well as in Chinese Meishan pigs showed that ART3 is mostly expressed in muscles enriched in fast  
389 twitch fibers<sup>47,48</sup>, which are typically more prone to damage in DMD. ART3 has been reported to be  
390 decreased in the serum across multiple dystrophies and myopathies<sup>39</sup> and in SBMA<sup>45</sup>. It has been  
391 recently reported in SBMA only 3 proteins, including RGMA, MSTN, and ART3 were found to have  
392 decreased abundances in plasma and were significantly correlated with higher thigh MRI muscle fat  
393 fraction akin to what is observed in DMD<sup>45</sup>. Additional support for the relationship to fat deposition  
394 has been obtained in Wannanhua pigs show that ART3 is involved in fat deposition in muscle<sup>49</sup>,  
395 further strengthening the biological rationale behind the association identified in this study. This is  
396 the first study to show ART3 relationship to CS status and clinical function in DMD.

397 Finally, EHMT2, known as euchromatic histone lysine methyltransferase 2, and DLK1 known as Delta-  
398 like 1 homolog as well as Pref-1 (preadipocyte factor 1) showed interesting association. They both  
399 declined with age and were associated with disease milestone; EHMT2 was normalized by steroid  
400 treatment. EHMT2 was previously associated with renal fibrosis<sup>50</sup>, atrial fibrosis<sup>51</sup>, cardiomyocytes

401 hypertrophy<sup>52</sup> and high fat diet induced obesity and hepatic insulin resistance<sup>53</sup>, which align with the  
402 pathogenic processes ongoing in DMD. The reduction of EHMT2 with age could potentially describe  
403 the reduced magnitude of pathological processes as muscle mass is progressively lost. DLK1 plays a  
404 multifaceted role in muscle development and regeneration. DLK1 is a transmembrane protein that  
405 functions as a regulator of cell growth during development. In adults, its expression is low and  
406 restricted to endocrine tissues. It is involved in the differentiation of multiple cell types, including  
407 adipocytes and plays an important role in skeletal muscle biology during fetal development and  
408 postnatal growth<sup>54</sup>. While the role of DLK1 in adult skeletal muscle regeneration is less clear,  
409 upregulated expression of DLK1 has previously been observed in DMD and Becker muscular  
410 dystrophy<sup>55</sup> and reduction of DLK1 in fibroadipogenic progenitors corresponded to increased  
411 adipogenic commitment<sup>56</sup>. The reduction of the DLK1 levels may be associated with an increased  
412 adipogenic signature in DMD.

413 One limitation of this study was the differences in cohort characteristics due to the participant pool  
414 and standards of care. The LUMC samples consisted of participants seen in the clinic, were generally  
415 younger in age, and were primarily undergoing a 10 days on, 10 days off CS regimen. No data were  
416 available on whether a patient was in the “on” or “off” phase of treatment at the time of sample  
417 collection. In contrast, the UF cohort samples were from participants in a natural history research  
418 study, were generally older in age, and were primarily on a daily CS regimen. The differences in  
419 steroid regimens may have led to smaller effect sizes associated with CS use in the LUMC cohort as  
420 well as the earlier occurrence of clinical milestones. The smaller proportion of individuals reaching  
421 milestones in the UF cohort may explain the smaller number of probes associated with clinical  
422 outcomes and milestones in that cohort. As different regimes coincided with cohort effects, we were  
423 unable to study the impact of intermittent CS compared to daily CS. These differences could  
424 potentially be studied in prospective research such as in the FOR-DMD clinical trial. Lastly, due to the  
425 retrospective nature of this study, we had a number of missing data points. Despite these  
426 limitations, due to our relatively large patient population and number of samples, we were able to  
427 identify multiple significant serum biomarkers which were validated across both cohorts.

428 In conclusion, we found proteins associated with upper and/or lower limb outcomes in individuals  
429 with DMD across two independent cohorts. Significant consideration should be given to the use of  
430 RGMA, DLK1, ANTXR2, EHMT2 and ART3 as potential biomarkers based on the strength of their  
431 associations, significance of the findings across scales and milestones and biological plausibility in  
432 connection to the disease. A panel that accurately and precisely detects these proteins in serum  
433 could enable the connection of short-term changes to disease stabilization and a decreased risk of  
434 decline in the mid- to long-term. Furthermore, measuring these proteins in a clinical setting would  
435 help predict individual disease trajectories, assess treatment options before milestones, and inform  
436 patient inclusion criteria in clinical trials. These findings are of importance for future clinical trial  
437 design, paving the way for innovative outcome measures such as serum biomarkers.

438

439

440 **Acknowledgements**

441 This study was funded by Parent Project Muscular Dystrophy through the Protein Mapping Project.  
442 Research reported in this publication was also partially supported by Spieren voor Spieren (grant  
443 number Svs15) and National Institute Of Neurological Disorders And Stroke of the National Institutes  
444 of Health under Award Number # R61NS119639 (Co-I: Spitali), R01 AR056973 (PI: K Vandenborne),  
445 and R01 AR065943 (PI: GA Walter). Research reported in this publication was also supported by the  
446 CTSI Biorepository at the University of Florida Clinical and Translational Science Institute, which is  
447 supported in part by the NIH National Center for Advancing Translational Sciences under award  
448 number UL1 TR001427. The authors thank the staff and participants for their contributions.

449

450 **Potential Conflicts of Interest**

451 SWME, SVH, and KCHH are employees of BioSymetrics, which has a commercial interest in the  
452 results. The remaining authors have no competing interests.

453

454 **References**

- 455 1. Mercuri, E. & Muntoni, F. Muscular dystrophies. *Lancet (London, England)* **381**, 845–60  
456 (2013).
- 457 2. Duan, D., Goemans, N., Takeda, S., Mercuri, E. & Aartsma-Rus, A. Duchenne muscular  
458 dystrophy. *Nature Reviews Disease Primers* vol. 7 1–19 (2021).
- 459 3. Bello, L. *et al.* Prednisone/prednisolone and deflazacort regimens in the CINRG Duchenne  
460 Natural History Study. *Neurology* **85**, 1048–1055 (2015).
- 461 4. McDonald, C. M. *et al.* Long-term effects of glucocorticoids on function, quality of life, and  
462 survival in patients with Duchenne muscular dystrophy: a prospective cohort study. *Lancet*  
463 *(London, England)* **391**, 451–461 (2018).
- 464 5. Birnkrant, D. J. *et al.* Review Diagnosis and management of Duchenne muscular dystrophy,  
465 part 1: diagnosis, and neuromuscular, rehabilitation, endocrine, and gastrointestinal and  
466 nutritional management. *Lancet Neurol.* **17**, 251–267 (2018).
- 467 6. Conklin, L. S. *et al.* Serum biomarkers of glucocorticoid response and safety in anti-neutrophil  
468 cytoplasmic antibody-associated vasculitis and juvenile dermatomyositis. *Steroids* **140**, 159–  
469 166 (2018).
- 470 7. Hoy, S. M. Delandistrogene Moxeparvovec: First Approval. *Drugs* **83**, 1323–1329 (2023).
- 471 8. Aartsma-Rus, A. The Future of Exon Skipping for Duchenne Muscular Dystrophy. *Hum. Gene*  
472 *Ther.* **34**, 372–378 (2023).
- 473 9. Keam, S. J. Vamorolone: First Approval. *Drugs* **84**, 111–117 (2024).
- 474 10. FDA Approves Nonsteroidal Treatment for Duchenne Muscular Dystrophy | FDA.  
475 [https://www.fda.gov/news-events/press-announcements/fda-approves-nonsteroidal-](https://www.fda.gov/news-events/press-announcements/fda-approves-nonsteroidal-treatment-duchenne-muscular-dystrophy)  
476 [treatment-duchenne-muscular-dystrophy.](https://www.fda.gov/news-events/press-announcements/fda-approves-nonsteroidal-treatment-duchenne-muscular-dystrophy)
- 477 11. Aartsma-Rus, A., Ginjaar, I. B. & Bushby, K. The importance of genetic diagnosis for Duchenne  
478 muscular dystrophy. *J. Med. Genet.* **53**, 145–151 (2016).
- 479 12. Wang, R. T. *et al.* DMD genotype correlations from the Duchenne Registry: Endogenous exon  
480 skipping is a factor in prolonged ambulation for individuals with a defined mutation subtype.  
481 *Hum. Mutat.* **39**, 1193–1202 (2018).
- 482 13. Bello, L. *et al.* DMD genotypes and loss of ambulation in the CINRG Duchenne Natural History  
483 Study. *Neurology* **87**, 401–409 (2016).
- 484 14. Waldrop, M. A. *et al.* Intron mutations and early transcription termination in Duchenne and  
485 Becker muscular dystrophy. *Hum. Mutat.* **43**, 511–528 (2022).
- 486 15. Bello, L. & Pegoraro, E. The “Usual Suspects”: Genes for Inflammation, Fibrosis, Regeneration,  
487 and Muscle Strength Modify Duchenne Muscular Dystrophy. *J. Clin. Med.* **8**, 649 (2019).
- 488 16. Alfano, L. *et al.* Role of motivation on performance of the 6-minute walk test in boys with  
489 Duchenne muscular dystrophy. *Dev. Med. Child Neurol.* **57**, 57–58 (2015).
- 490 17. Al-Khalili Szigyarto, C. & Spitali, P. Biomarkers of Duchenne muscular dystrophy: current  
491 findings. *Degener. Neurol. Neuromuscul. Dis.* **Volume 8**, 1–13 (2018).

- 492 18. Molinaro, M., Torrente, Y., Villa, C. & Farini, A. Advancing Biomarker Discovery and  
493 Therapeutic Targets in Duchenne Muscular Dystrophy: A Comprehensive Review. *Int. J. Mol.*  
494 *Sci.* **25**, (2024).
- 495 19. Hathout, Y. *et al.* Disease-specific and glucocorticoid-responsive serum biomarkers for  
496 Duchenne Muscular Dystrophy. *Sci. Rep.* **9**, 12167 (2019).
- 497 20. Wagner, K. R. *et al.* Safety and disease monitoring biomarkers in Duchenne muscular  
498 dystrophy: Results from a Phase II trial. *Biomark. Med.* **15**, 1389–1396 (2021).
- 499 21. Zygmont, A. M. *et al.* A longitudinal study of creatine kinase and creatinine levels in  
500 Duchenne muscular dystrophy. *Muscle Nerve* **67**, 138–145 (2023).
- 501 22. Awano, H. *et al.* Longitudinal data of serum creatine kinase levels and motor, pulmonary, and  
502 cardiac functions in 337 patients with Duchenne muscular dystrophy. *Muscle Nerve* **69**, 604–  
503 612 (2024).
- 504 23. Boehler, J. F. *et al.* Clinical potential of microdystrophin as a surrogate endpoint.  
505 *Neuromuscul. Disord.* **33**, 40–49 (2023).
- 506 24. Chamberlain, J. S. *et al.* Microdystrophin Expression as a Surrogate Endpoint for Duchenne  
507 Muscular Dystrophy Clinical Trials. *Hum. Gene Ther.* **34**, 404 (2023).
- 508 25. Spitali, P. *et al.* Tracking disease progression non-invasively in Duchenne and Becker muscular  
509 dystrophies. *J. Cachexia. Sarcopenia Muscle* **9**, 715–726 (2018).
- 510 26. Strandberg, K. *et al.* Blood-derived biomarkers correlate with clinical progression in Duchenne  
511 muscular dystrophy. *J. Neuromuscul. Dis.* **7**, 231–246 (2020).
- 512 27. Scott, E. *et al.* Development of a functional assessment scale for ambulatory boys with  
513 Duchenne muscular dystrophy. *Physiother. Res. Int.* **17**, 101–109 (2012).
- 514 28. Arora, H. *et al.* Longitudinal timed function tests in Duchenne muscular dystrophy:  
515 ImagingDMD cohort natural history. *Muscle Nerve* **58**, 631–638 (2018).
- 516 29. McDonald, C. M. *et al.* The 6-minute walk test and other clinical endpoints in duchenne  
517 muscular dystrophy: Reliability, concurrent validity, and minimal clinically important  
518 differences from a multicenter study. *Muscle Nerve* **48**, 357–368 (2013).
- 519 30. Pane, M. *et al.* Upper limb function in Duchenne muscular dystrophy: 24 month longitudinal  
520 data. *PLoS One* **13**, (2018).
- 521 31. Candia, J., Daya, G. N., Tanaka, T., Ferrucci, L. & Walker, K. A. Assessment of variability in the  
522 plasma 7k SomaScan proteomics assay. *Sci. Rep.* **12**, (2022).
- 523 32. Jolly, E. Pymer4: Connecting R and Python for Linear Mixed Modeling. *J. Open Source Softw.*  
524 **3**, 862 (2018).
- 525 33. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using  
526 lme4. *J. Stat. Softw.* **67**, (2015).
- 527 34. Davidson-Pilon, C. lifelines: survival analysis in Python. *J. Open Source Softw.* **4**, 1317 (2019).
- 528 35. Signorelli, M. *et al.* Longitudinal serum biomarker screening identifies malate dehydrogenase  
529 2 as candidate prognostic biomarker for Duchenne muscular dystrophy. *J. Cachexia.*  
530 *Sarcopenia Muscle* **11**, 505–517 (2020).



- 531 36. Hathout, Y. *et al.* Large-scale serum protein biomarker discovery in Duchenne muscular  
532 dystrophy. *Proc. Natl. Acad. Sci.* **112**, 7153–7158 (2015).
- 533 37. Alonso-Jiménez, A. *et al.* Platelet Derived Growth Factor-AA Correlates With Muscle Function  
534 Tests and Quantitative Muscle Magnetic Resonance in Dystrophinopathies. *Front. Neurol.* **12**,  
535 659922 (2021).
- 536 38. Hathout, Y. *et al.* Serum pharmacodynamic biomarkers for chronic corticosteroid treatment  
537 of children. *Sci. Rep.* **6**, 31727 (2016).
- 538 39. Willis, A. B. *et al.* Serum protein and imaging biomarkers after intermittent steroid treatment  
539 in muscular dystrophy. 1–24 (2024) doi:10.1101/2024.06.14.24308858.
- 540 40. Milad, N. *et al.* Increased plasma lipid levels exacerbate muscle pathology in the mdx mouse  
541 model of Duchenne muscular dystrophy. *Skelet. Muscle* **7**, 19 (2017).
- 542 41. White, Z. *et al.* Cholesterol absorption blocker ezetimibe prevents muscle wasting in severe  
543 dysferlin-deficient and mdx mice. *J. Cachexia. Sarcopenia Muscle* **13**, 544–560 (2022).
- 544 42. do Carmo Costa, A. *et al.* RGMa can induce skeletal muscle cell hyperplasia via association  
545 with neogenin signalling pathway. *Vitr. Cell. Dev. Biol. - Anim.* **57**, 415–427 (2021).
- 546 43. Copola, A. G. L. *et al.* Transcriptomic characterization of the molecular mechanisms induced  
547 by RGMa during skeletal muscle nuclei accretion and hypertrophy. *BMC Genomics* **23**, (2022).
- 548 44. Neto, J. X. *et al.* Repulsive Guidance Molecules a, b and c Are Skeletal Muscle Proteins, and  
549 Repulsive Guidance Molecule a Promotes Cellular Hypertrophy and Is Necessary for Myotube  
550 Fusion Aline Fagundes Martins<sup>1</sup>. *Cells Tissues Organs* **200**, 326–338 (2015).
- 551 45. Tebbenkamp, A. T. N. *et al.* Protein biomarker signature in patients with spinal and bulbar  
552 muscular atrophy. *JCI Insight* **9**, (2024).
- 553 46. Bürgi, J. *et al.* CMG2/ANTXR2 regulates extracellular collagen VI which accumulates in hyaline  
554 fibromatosis syndrome. *Nat. Commun.* 2017 81 **8**, 1–10 (2017).
- 555 47. Abbassi-Daloui, T. *et al.* A transcriptome atlas of leg muscles from healthy human volunteers  
556 reveals molecular and cellular signatures associated with muscle location. *Elife* **12**, 1–29  
557 (2023).
- 558 48. Li, Y., Xu, Z., Li, H., Xiong, Y. & Zuo, B. Differential transcriptional analysis between red and  
559 white skeletal muscle of Chinese Meishan pigs. *Int. J. Biol. Sci.* **6**, 350–360 (2010).
- 560 49. Li, X. *et al.* Integrated 4D Analysis of Intramuscular Fat Deposition: Quantitative Proteomic  
561 and Transcriptomic Studies in Wannanhua Pig Longissimus Dorsi Muscle. *Animals* **14**, 167  
562 (2024).
- 563 50. Irifuku, T. *et al.* Inhibition of H3K9 histone methyltransferase G9a attenuates renal fibrosis  
564 and retains klotho expression. *Kidney Int.* **89**, 147–157 (2016).
- 565 51. Xiao, Z. *et al.* MicroRNA-205-5p plays a suppressive role in the high-fat diet-induced atrial  
566 fibrosis through regulation of the EHMT2/IGFBP3 axis. *Genes Nutr.* **17**, 11 (2022).
- 567 52. Papait, R. *et al.* Histone Methyltransferase G9a Is Required for Cardiomyocyte Homeostasis  
568 and Hypertrophy. *Circulation* **136**, 1233–1246 (2017).
- 569 53. Xue, W. *et al.* Histone methyltransferase G9a modulates hepatic insulin signaling via

- 570 regulating HMGA1. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1864**, 338–346 (2018).
- 571 54. Zhang, L. *et al.* Expression and Functional Analyses of Dlk1 in Muscle Stem Cells and  
572 Mesenchymal Progenitors during Muscle Regeneration. *Int. J. Mol. Sci.* **20**, (2019).
- 573 55. Andersen, D. C. *et al.* Characterization of DLK1+ cells emerging during skeletal muscle  
574 remodeling in response to myositis, myopathies, and acute injury. *Stem Cells* **27**, 898–908  
575 (2009).
- 576 56. Garcia, S. M. *et al.* Distinct human stem cell subpopulations drive adipogenesis and fibrosis in  
577 musculoskeletal injury. *bioRxiv Prepr. Serv. Biol.* (2024) doi:10.1101/2023.07.28.551038.
- 578
- 579
- 580

581 **Figure legends**

582 **Figure 1.** Associations of Serum Proteins with Age in Individuals with DMD. Volcano plots  
583 demonstrate the proteins from DMD serum that are associated with age in A) the LUMC cohort, and  
584 B) the UF cohort. C) Venn diagram showing the number of proteins associated with age in common  
585 between the LUMC and UF cohorts, with D) a plot of coefficients, indicating the direction of protein  
586 association with age, in the LUMC versus UF cohorts. Plots E-H highlight the longitudinal trajectories  
587 of protein levels over time in select proteins that decrease with increasing age such as CKM,  
588 MYOM3, TTN and TNNI2, while plots I-L show trajectories of select proteins that increase with  
589 increasing age such as CNTN3, CNDP1, LEP and TIMP4. Each trajectory plot includes the gene symbol  
590 and aptamer identification number. DMD, Duchenne muscular dystrophy; LUMC, Leiden University  
591 Medical Center; UF, University of Florida.

592 **Figure 2.** Associations of Serum Proteins with Corticosteroid use in Individuals with DMD. Volcano  
593 plots demonstrate the proteins from DMD serum that are associated with corticosteroid use in A)  
594 the LUMC cohort, and B) the UF cohort. C) Venn diagram showing the number of proteins associated  
595 with corticosteroid use in common between the LUMC and UF cohorts, with D) a plot of coefficients,  
596 indicating the direction of protein association with corticosteroid use, in the LUMC versus UF  
597 cohorts. Plots E-H show differences in protein levels between serum from CS treated versus  
598 untreated individuals with DMD in two previously identified proteins and two proteins identified in  
599 this study. I) Scatterplot showing the relationship between age and steroid treatment coefficients for  
600 proteins significantly associated with CS treatment. Proteins showing discordant coefficients for age  
601 and CS treatments were considered as efficacy biomarkers, while with discordant coefficients were  
602 considered as safety biomarkers (grey shaded). Orange dots represent estimates for the LUMC  
603 cohort, while blue dots represent estimates for the UF cohort. CS, corticosteroid; DMD, Duchenne  
604 muscular dystrophy; LUMC, Leiden University Medical Center; UF, University of Florida.

605 **Figure 3.** Associations of Serum Proteins with Performance on Tests of Motor Function in Individuals  
606 with DMD. A-D) Longitudinal trajectory plots of NSAA results, 10MRV, 6MWD, and PUL 2.0 score  
607 from both cohorts. E) Plots of coefficients, indicating the direction of protein association with motor  
608 function test performance in the LUMC versus UF cohorts. Highlighted data points represent  
609 significant associations. The number of significant associations is mentioned for each panel. F) UpSet  
610 plot demonstrating the intersection of significantly associated probes with each of the four tests of  
611 motor function. G-H) Trajectory plots showing the relationship between ART3 and either NSAA (G) or  
612 6MWT (H) over the age in the LUMC cohort. I) Scatterplot showing the relationship between the  
613 coefficient estimates for proteins significantly associated with steroid treatment and performance  
614 test. Proteins showing concordant coefficients for performance tests and CS treatments were  
615 considered as efficacy biomarkers (e.g. a protein associated with both high functional scores and  
616 further elevated by steroid treatment), while with concordant coefficients were considered as safety  
617 biomarkers (grey shaded). Orange dots represent estimates for the LUMC cohort, while blue dots  
618 represent estimates for the UF cohort. Shapes indicate the association for each performance test.  
619 6MWD, 6-minute walk distance; 10MRV, 10-meter run velocity; CS, corticosteroid; DMD, Duchenne  
620 muscular dystrophy; LUMC, Leiden University Medical Center; NSAA, North Star Ambulatory  
621 Assessment; PUL, Performance of the Upper Limb; UF, University of Florida.

622

623 **Figure 4.** Associations of Serum Proteins with Clinical Milestones. A) Kaplan Meier curves showing  
624 the probability of LoA, loss of OHR, and loss of HTM by age in each cohort. B) UpSet plot of the  
625 intersection of significant probes with each clinical milestone. C) Log hazard ratios of the protein  
626 probes significantly associated with at least 1 clinical milestone. D) Kaplan Meier curves for loss of  
627 each clinical milestone for RGMA by years since first sample collection. E) Correlations between  
628 protein probe log hazard ratios and use of corticosteroids. HTM, hand-to-mouth; LoA, loss of  
629 ambulation; OHR, overhead reach.

630

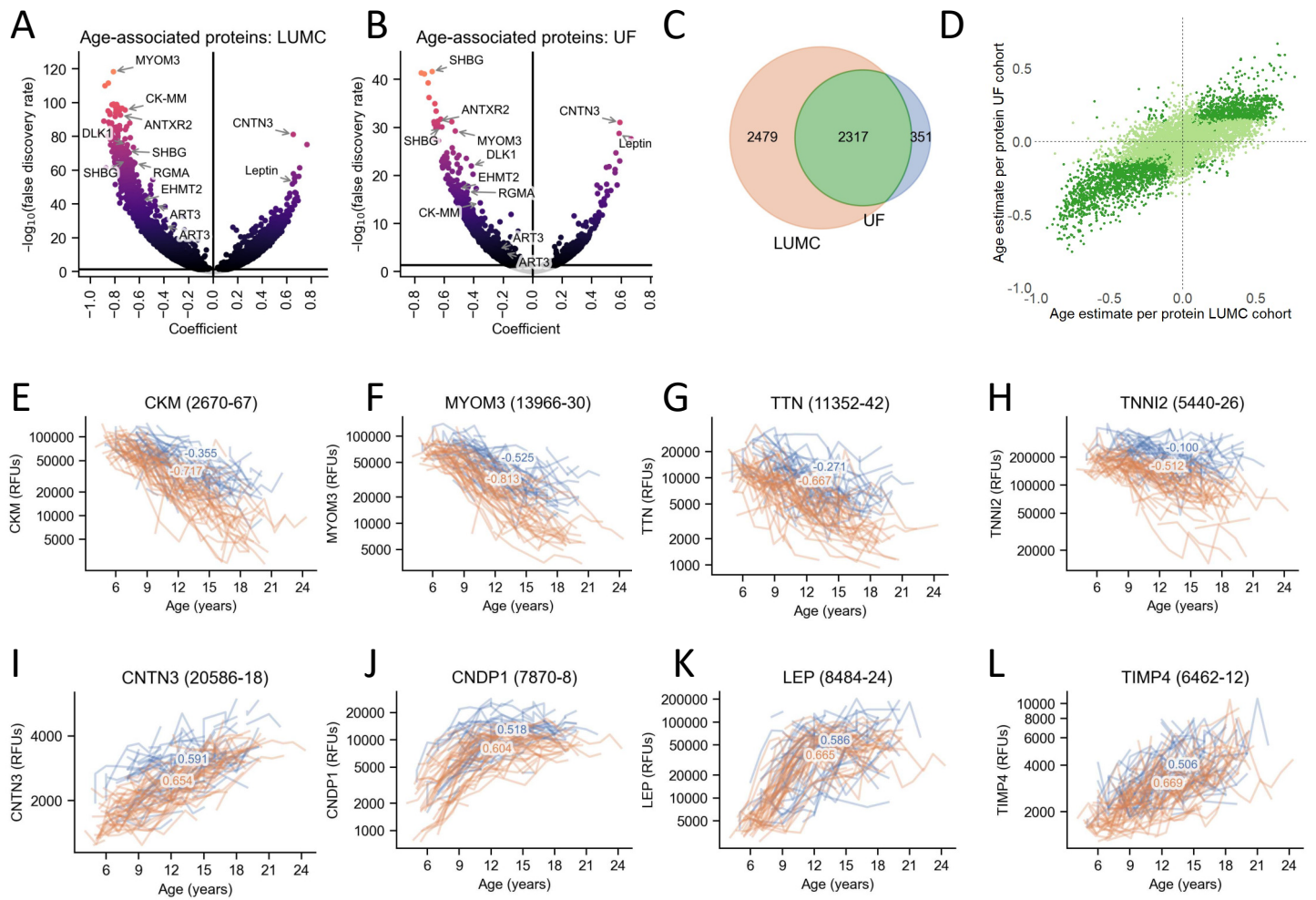


Figure 1

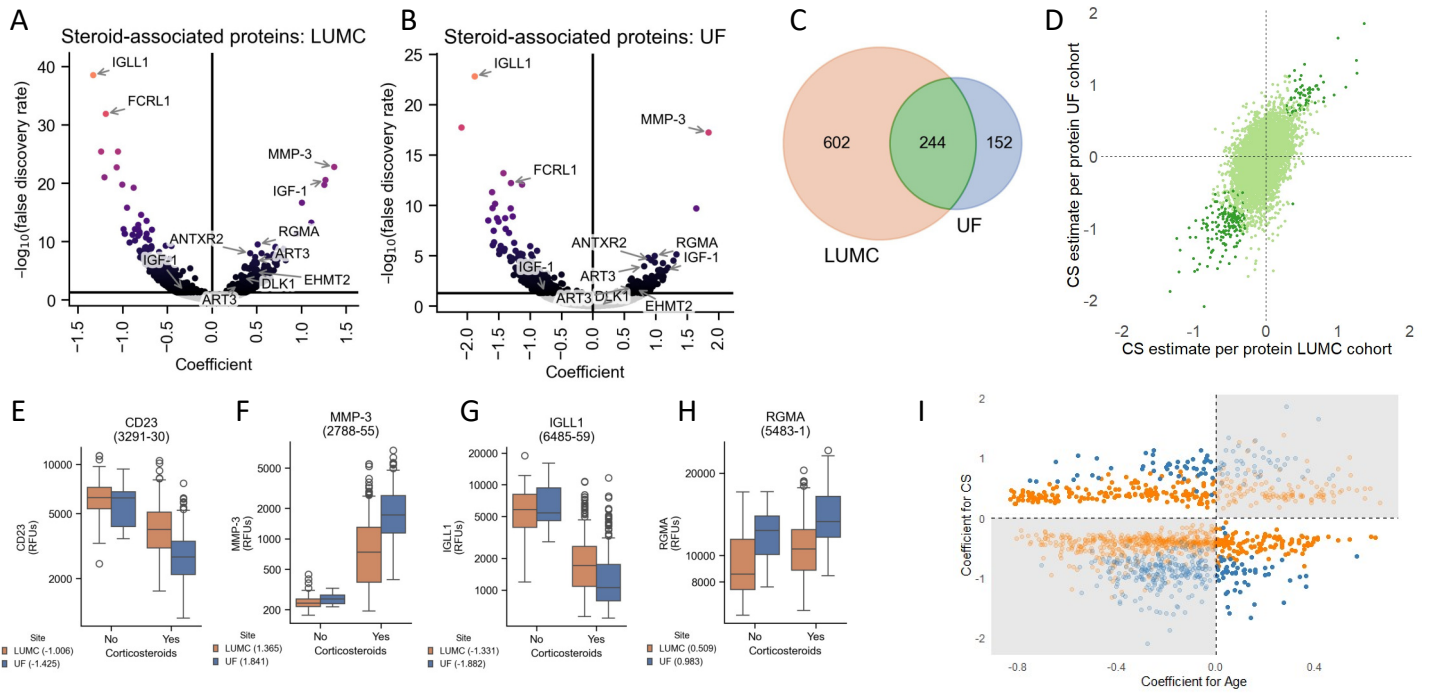


Figure 2

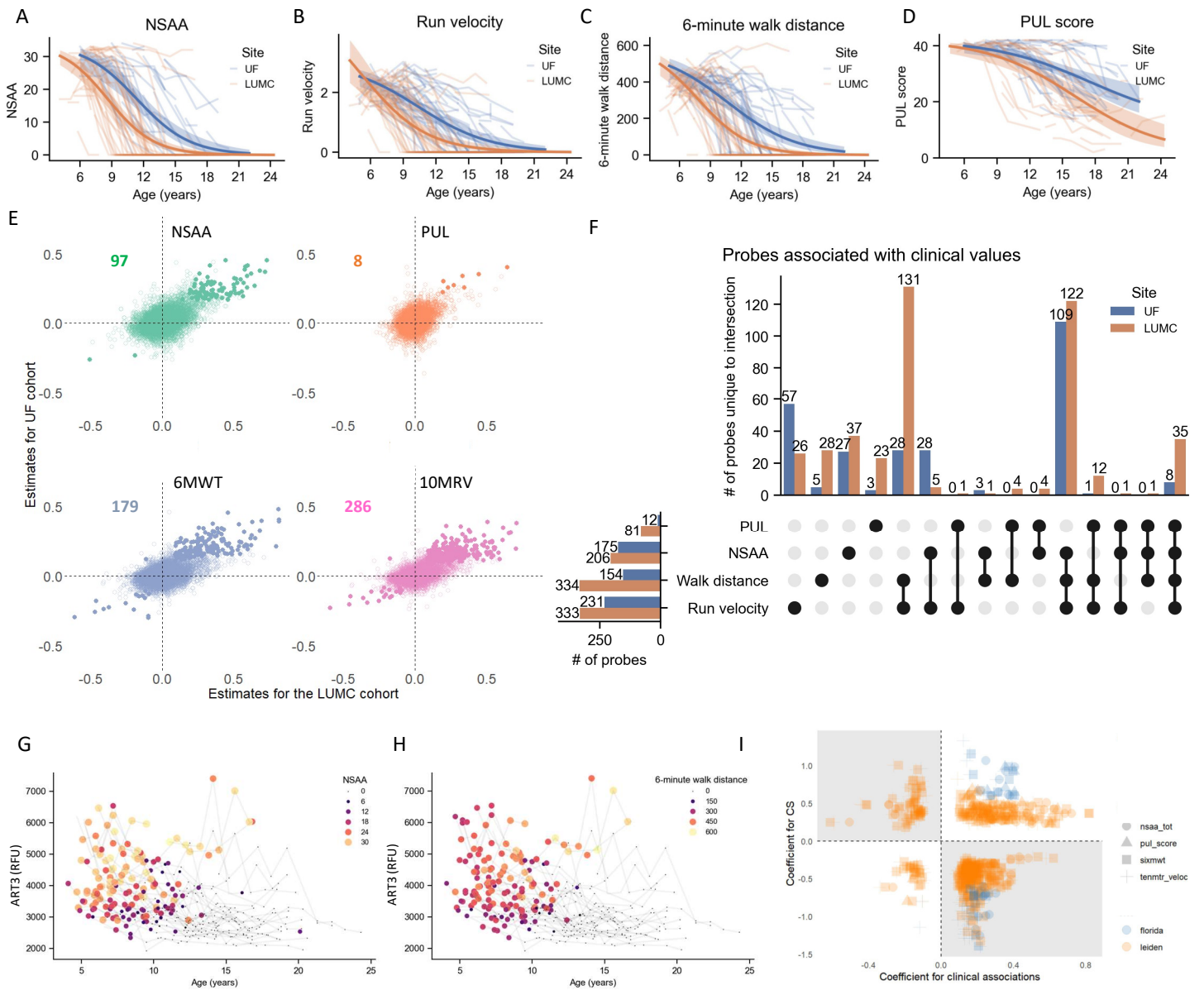


Figure 3

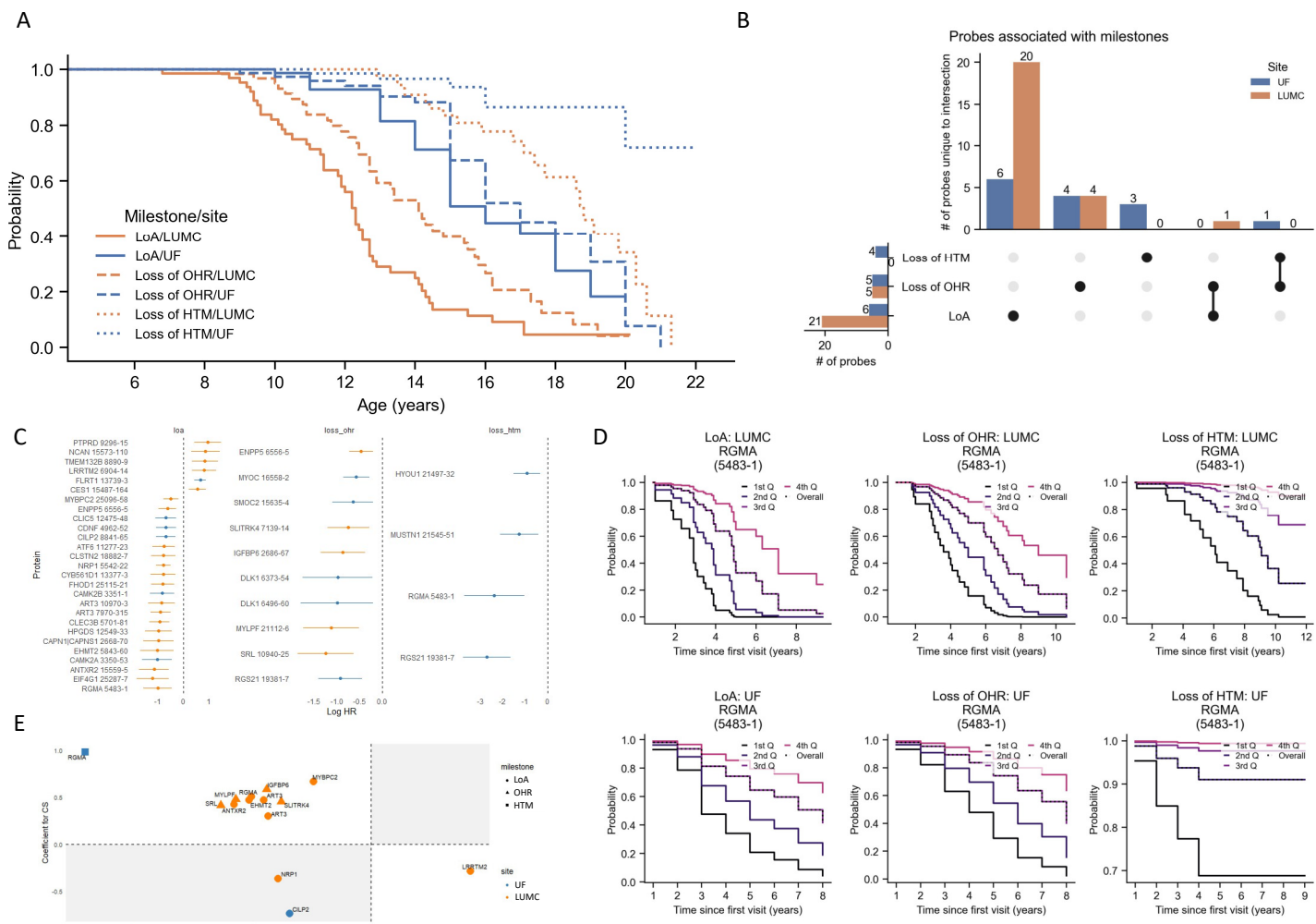


Figure 4