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

Authors:

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

Affiliations:
1. Department of Neurology, Leiden University Medical Center, Leiden, Zuid-Holland, the Netherlands
2. Duchenne Center Netherlands
3. Department of Physical Therapy, University of Florida, Gainesville, FL, USA
4. BioSymetrics, Inc., Huntington, NY, USA.
5. C.J. Gorter MRI Center, Department of Radiology, Leiden University Medical Center, Leiden, Zuid-Holland, the Netherlands
6. Department of Physiology and Aging, University of Florida, Gainesville, FL, USA
7. Human Genetics Department, Leiden University Medical Center, Leiden, Zuid-Holland, the Netherlands

* Authors contributed equally

§ Authors contributed equally

Abstract

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

Introduction

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

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

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

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

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

Methods

Study Cohort, design and outcomes

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

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

Samples collection and proteomic analysis

Serum samples were collected and prepared according to standard phlebotomy procedures. Samples at LUMC were left to clot for ~30 minutes, followed by 10 min centrifugation at 2350g, while clotting time was 30 minutes at UF, followed by centrifugation at 3000 rpm for 15 min. At the LUMC, sample aliquots were frozen at -20°C for 1-2- months and then transferred to -80°C for long term storage, while at UF, samples were immediately stored at -80°C. Serum samples were analysed by the SomaScan® proteomic platform at SomaLogic (Boulder, Colorado, USA)[31]. SomaScan uses aptamers as affinity reagents to detect and quantify proteins in complex mixtures such as serum. Aptamers are nucleic acids that are able to bind specific protein targets with high specificity and affinity. The pipeline is built to keep targets engaging aptamers, while washing away the unbound ones. Detection of the aptamers is then performed using an array with probes complementary to the aptamer sequence. At the time of this research, this platform included 7596 aptamers capable of detecting 6628 proteins. As part of the assay, extensive quality control metrics are put in place including internal and external controls. A total of 9 samples (2 LUMC / 7 UF) did not pass SomaLogic's internal quality control and were excluded from further analysis, results, and data 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 *pymer4* package, version 0.7.8, a Python interface for the *lme4* R
169 package, version 1.1-31[32,33]. 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 95th 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[34]. 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

Results

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

	LUMC n = 74	**UF n = 79**
Samples, n	405	288
Age at first sample, years, mean (SD)	8.4 (3.4)	10.9 (3.2)
Follow-up duration, years, mean (SD)	5.7 (3.6)	3.4 (2.5)
Samples per patient, n, mean (SD)	5.5 (2.9)	3.6 (2.0)
Age at start of corticosteroids, years, mean (SD)	5.6 (1.6)	5.4 (1.7)
Corticosteroid treatment at first sample, n (%)		
• *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%)
Dystrophin variants, n (%)		
• *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%)
Weight at first sample, kg, mean (SD)	32.6 (18.6)	35.1 (12.2)
Height at first sample, cm, mean (SD)	127.2 (19.2)	126.9 (10.7)
BMI at first sample, mean (SD)	18.3 (4.4)	21.2 (5.4)
Ambulant patients at first sample, n (%)	59 (79.7%)	68/78 (87.2%) (1 unknown)
Able to perform overhead reach at first sample, n (%)	64 (86.5%)	63/70 (90.0%) (9 unknown)
Able to perform hand to mouth at first sample, n (%)	73 (98.6%)	74/75 (98.7%) (4 unknown)

Table 1. Cohort characteristics.

Serum protein associations with age and corticosteroid usage

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

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

Among the proteins with expression that significantly increased with age, we identified proteins present in the central nervous system such as CNTN3 and CNDP1, proteins expressed in adipose tissue (LEP, GHR, FABP4, ADIPOQ, CNTFR, CFD and TIMP4, PNLIPRP1, PNLIPRP2, PNLIP), and several cytokines (CCL14, IFNA6, BMPR2, CXCL8, CCL3L1, IL26, IL5RA, CCL1, CCL16, CCL15, IL10RA, IL31RA, IL17RE) (Fig1I-L). Two different probes for IGF1 had had opposing associations with age, and several IGF binding proteins showed discordant trajectories, with IGFBP1 and IGFBP2 decreasing with age andIGFBP5 and IGFBP6 increasing with age. Coefficients of all significant, concordant proteins associated with age are provided in Supplementary Table 1.

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

<u>Serum protein associations with motor performance</u>

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

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

Prediction of clinical milestones

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

Discussion

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

Prior serum biomarker discovery in dystrophinopathies has largely been limited to smaller cohorts, targeted protein selection for analysis, and limited longitudinal data. These studies have been crucial for building our understanding of the disease's pathogenic processes. Key differences between individuals with and without DMD have been reported, including several proteins associated with age, shedding light on the ongoing pathogenic processes. It is well-known with DMD that serum CK (activity and protein levels) decrease with age, reflecting muscle damage and loss of muscle mass. Additional proteins decreasing with age have subsequently been identified including MDH2, MYL3, CA3, MYOM3, COL1A1, TTN, TNNT2, ETFA, TNNT3, and MAP4[26,35]. Proteins that have increasing expression with age have included C4A, C4BPA, and GSN, among others[35]. A small number of studies have also utilized the SomaScan® platform to perform large-scale protein biomarker analysis in DMD, identifying additional proteins associated with age/time such as LEP, CFTNR, FABP3, and TNNI2[25,36]. Complementing and largely validating the findings of these studies, we identified previously described proteins such as CK and MYOM-3 showing steep decline in protein levels, but also substitution of muscle with fibro-adipogenic tissue with markers such as Collagen 1 and 2 declining along with markers of fibro-adipogenic progenitors such as PDGFRA[37]. We also observed adipogenic markers such as LEP, GHR and ADIPOQ with increased expression along with a complement and inflammation signature.

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

Treatment with corticosteroids contributed to normalize as well as exacerbate the disease progression signature. An exacerbation of the signature related to age was observed for previously reported safety biomarkers such as MMP3, Afamin, and IGFBP5[6], and for certain apolipoproteins such as APOA2, APOL1, APOA5. A normalization was observed for other apolipoproteins such as APOE4, APOC3 and APOE, suggesting that steroid treatment could directly normalize the dyslipidemia with effects on APOE4, APOC3, and APOE, while exacerbation of the APOA2, APOL1, and APOA5 signature could highlight how steroids affect lipid metabolism and potentially

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

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

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

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

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

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

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

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

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

Acknowledgements

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

Potential Conflicts of Interest

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

References

1. Mercuri, E. & Muntoni, F. Muscular dystrophies. *Lancet (London, England)* **381**, 845–60 (2013).

2. Duan, D., Goemans, N., Takeda, S., Mercuri, E. & Aartsma-Rus, A. Duchenne muscular dystrophy. *Nature Reviews Disease Primers* vol. 7 1–19 (2021).

3. Bello, L. *et al.* Prednisone/prednisolone and deflazacort regimens in the CINRG Duchenne Natural History Study. *Neurology* **85**, 1048–1055 (2015).

4. McDonald, C. M. *et al.* Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study. *Lancet (London, England)* **391**, 451–461 (2018).

5. Birnkrant, D. J. *et al.* Review Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and neuromuscular, rehabilitation, endocrine, and gastrointestinal and nutritional management. *Lancet Neurol.* **17**, 251–267 (2018).

6. Conklin, L. S. *et al.* Serum biomarkers of glucocorticoid response and safety in anti-neutrophil cytoplasmic antibody-associated vasculitis and juvenile dermatomyositis. *Steroids* **140**, 159–166 (2018).

7. Hoy, S. M. Delandistrogene Moxeparvovec: First Approval. *Drugs* **83**, 1323–1329 (2023).

8. Aartsma-Rus, A. The Future of Exon Skipping for Duchenne Muscular Dystrophy. *Hum. Gene Ther.* **34**, 372–378 (2023).

9. Keam, S. J. Vamorolone: First Approval. *Drugs* **84**, 111–117 (2024).

10. FDA Approves Nonsteroidal Treatment for Duchenne Muscular Dystrophy | FDA. https://www.fda.gov/news-events/press-announcements/fda-approves-nonsteroidal-treatment-duchenne-muscular-dystrophy.

11. Aartsma-Rus, A., Ginjaar, I. B. & Bushby, K. The importance of genetic diagnosis for Duchenne muscular dystrophy. *J. Med. Genet.* **53**, 145–151 (2016).

12. Wang, R. T. *et al.* DMD genotype correlations from the Duchenne Registry: Endogenous exon skipping is a factor in prolonged ambulation for individuals with a defined mutation subtype. *Hum. Mutat.* **39**, 1193–1202 (2018).

13. Bello, L. *et al.* DMD genotypes and loss of ambulation in the CINRG Duchenne Natural History Study. *Neurology* **87**, 401–409 (2016).

14. Waldrop, M. A. *et al.* Intron mutations and early transcription termination in Duchenne and Becker muscular dystrophy. *Hum. Mutat.* **43**, 511–528 (2022).

15. Bello, L. & Pegoraro, E. The “Usual Suspects”: Genes for Inflammation, Fibrosis, Regeneration, and Muscle Strength Modify Duchenne Muscular Dystrophy. *J. Clin. Med.* **8**, 649 (2019).

16. Alfano, L. *et al.* Role of motivation on performance of the 6-minute walk test in boys with Duchenne muscular dystrophy. *Dev. Med. Child Neurol.* **57**, 57–58 (2015).

17. Al-Khalili Szigyarto, C. & Spitali, P. Biomarkers of Duchenne muscular dystrophy: current findings. *Degener. Neurol. Neuromuscul. Dis.* **Volume 8**, 1–13 (2018).

18. Molinaro, M., Torrente, Y., Villa, C. & Farini, A. Advancing Biomarker Discovery and Therapeutic Targets in Duchenne Muscular Dystrophy: A Comprehensive Review. *Int. J. Mol. Sci.* **25**, (2024).

19. Hathout, Y. *et al.* Disease-specific and glucocorticoid-responsive serum biomarkers for Duchenne Muscular Dystrophy. *Sci. Rep.* **9**, 12167 (2019).

20. Wagner, K. R. *et al.* Safety and disease monitoring biomarkers in Duchenne muscular dystrophy: Results from a Phase II trial. *Biomark. Med.* **15**, 1389–1396 (2021).

21. Zygmunt, A. M. *et al.* A longitudinal study of creatine kinase and creatinine levels in Duchenne muscular dystrophy. *Muscle Nerve* **67**, 138–145 (2023).

22. Awano, H. *et al.* Longitudinal data of serum creatine kinase levels and motor, pulmonary, and cardiac functions in 337 patients with Duchenne muscular dystrophy. *Muscle Nerve* **69**, 604–612 (2024).

23. Boehler, J. F. *et al.* Clinical potential of microdystrophin as a surrogate endpoint. *Neuromuscul. Disord.* **33**, 40–49 (2023).

24. Chamberlain, J. S. *et al.* Microdystrophin Expression as a Surrogate Endpoint for Duchenne Muscular Dystrophy Clinical Trials. *Hum. Gene Ther.* **34**, 404 (2023).

25. Spitali, P. *et al.* Tracking disease progression non-invasively in Duchenne and Becker muscular dystrophies. *J. Cachexia. Sarcopenia Muscle* **9**, 715–726 (2018).

26. Strandberg, K. *et al.* Blood-derived biomarkers correlate with clinical progression in Duchenne muscular dystrophy. *J. Neuromuscul. Dis.* **7**, 231–246 (2020).

27. Scott, E. *et al.* Development of a functional assessment scale for ambulatory boys with Duchenne muscular dystrophy. *Physiother. Res. Int.* **17**, 101–109 (2012).

28. Arora, H. *et al.* Longitudinal timed function tests in Duchenne muscular dystrophy: ImagingDMD cohort natural history. *Muscle Nerve* **58**, 631–638 (2018).

29. McDonald, C. M. *et al.* The 6-minute walk test and other clinical endpoints in duchenne muscular dystrophy: Reliability, concurrent validity, and minimal clinically important differences from a multicenter study. *Muscle Nerve* **48**, 357–368 (2013).

30. Pane, M. *et al.* Upper limb function in Duchenne muscular dystrophy: 24 month longitudinal data. *PLoS One* **13**, (2018).

31. Candia, J., Daya, G. N., Tanaka, T., Ferrucci, L. & Walker, K. A. Assessment of variability in the plasma 7k SomaScan proteomics assay. *Sci. Rep.* **12**, (2022).

32. Jolly, E. Pymer4: Connecting R and Python for Linear Mixed Modeling. *J. Open Source Softw.* **3**, 862 (2018).

33. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **67**, (2015).

34. Davidson-Pilon, C. lifelines: survival analysis in Python. *J. Open Source Softw.* **4**, 1317 (2019).

35. Signorelli, M. *et al.* Longitudinal serum biomarker screening identifies malate dehydrogenase 2 as candidate prognostic biomarker for Duchenne muscular dystrophy. *J. Cachexia. Sarcopenia Muscle* **11**, 505–517 (2020).

36. Hathout, Y. *et al.* Large-scale serum protein biomarker discovery in Duchenne muscular dystrophy. *Proc. Natl. Acad. Sci.* **112**, 7153–7158 (2015).

37. Alonso-Jiménez, A. *et al.* Platelet Derived Growth Factor-AA Correlates With Muscle Function Tests and Quantitative Muscle Magnetic Resonance in Dystrophinopathies. *Front. Neurol.* **12**, 659922 (2021).

38. Hathout, Y. *et al.* Serum pharmacodynamic biomarkers for chronic corticosteroid treatment of children. *Sci. Rep.* **6**, 31727 (2016).

39. Willis, A. B. *et al.* Serum protein and imaging biomarkers after intermittent steroid treatment in muscular dystrophy. 1–24 (2024) doi:10.1101/2024.06.14.24308858.

40. Milad, N. *et al.* Increased plasma lipid levels exacerbate muscle pathology in the mdx mouse model of Duchenne muscular dystrophy. *Skelet. Muscle* **7**, 19 (2017).

41. White, Z. *et al.* Cholesterol absorption blocker ezetimibe prevents muscle wasting in severe dysferlin-deficient and mdx mice. *J. Cachexia. Sarcopenia Muscle* **13**, 544–560 (2022).

42. do Carmo Costa, A. *et al.* RGMa can induce skeletal muscle cell hyperplasia via association with neogenin signalling pathway. *Vitr. Cell. Dev. Biol. - Anim.* **57**, 415–427 (2021).

43. Copola, A. G. L. *et al.* Transcriptomic characterization of the molecular mechanisms induced by RGMa during skeletal muscle nuclei accretion and hypertrophy. *BMC Genomics* **23**, (2022).

44. Neto, J. X. *et al.* Repulsive Guidance Molecules a, b and c Are Skeletal Muscle Proteins, and Repulsive Guidance Molecule a Promotes Cellular Hypertrophy and Is Necessary for Myotube Fusion Aline Fagundes Martins1. *Cells Tissues Organs* **200**, 326–338 (2015).

45. Tebbenkamp, A. T. N. *et al.* Protein biomarker signature in patients with spinal and bulbar muscular atrophy. *JCI Insight* **9**, (2024).

46. Bürgi, J. *et al.* CMG2/ANTXR2 regulates extracellular collagen VI which accumulates in hyaline fibromatosis syndrome. *Nat. Commun. 2017 81* **8**, 1–10 (2017).

47. Abbassi-Daloii, T. *et al.* A transcriptome atlas of leg muscles from healthy human volunteers reveals molecular and cellular signatures associated with muscle location. *Elife* **12**, 1–29 (2023).

48. Li, Y., Xu, Z., Li, H., Xiong, Y. & Zuo, B. Differential transcriptional analysis between red and white skeletal muscle of Chinese Meishan pigs. *Int. J. Biol. Sci.* **6**, 350–360 (2010).

49. Li, X. *et al.* Integrated 4D Analysis of Intramuscular Fat Deposition: Quantitative Proteomic and Transcriptomic Studies in Wannanhua Pig Longissimus Dorsi Muscle. *Animals* **14**, 167 (2024).

50. Irifuku, T. *et al.* Inhibition of H3K9 histone methyltransferase G9a attenuates renal fibrosis and retains klotho expression. *Kidney Int.* **89**, 147–157 (2016).

51. Xiao, Z. *et al.* MicroRNA-205-5p plays a suppressive role in the high-fat diet-induced atrial fibrosis through regulation of the EHMT2/IGFBP3 axis. *Genes Nutr.* **17**, 11 (2022).

52. Papait, R. *et al.* Histone Methyltransferase G9a Is Required for Cardiomyocyte Homeostasis and Hypertrophy. *Circulation* **136**, 1233–1246 (2017).

53. Xue, W. *et al.* Histone methyltransferase G9a modulates hepatic insulin signaling via

regulating HMGA1. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1864**, 338–346 (2018).

54. Zhang, L. *et al.* Expression and Functional Analyses of Dlk1 in Muscle Stem Cells and Mesenchymal Progenitors during Muscle Regeneration. *Int. J. Mol. Sci.* **20**, (2019).

55. Andersen, D. C. *et al.* Characterization of DLK1+ cells emerging during skeletal muscle remodeling in response to myositis, myopathies, and acute injury. *Stem Cells* **27**, 898–908 (2009).

56. Garcia, S. M. *et al.* Distinct human stem cell subpopulations drive adipogenesis and fibrosis in musculoskeletal injury. *bioRxiv Prepr. Serv. Biol.* (2024) doi:10.1101/2023.07.28.551038.

Figure legends

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

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

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

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

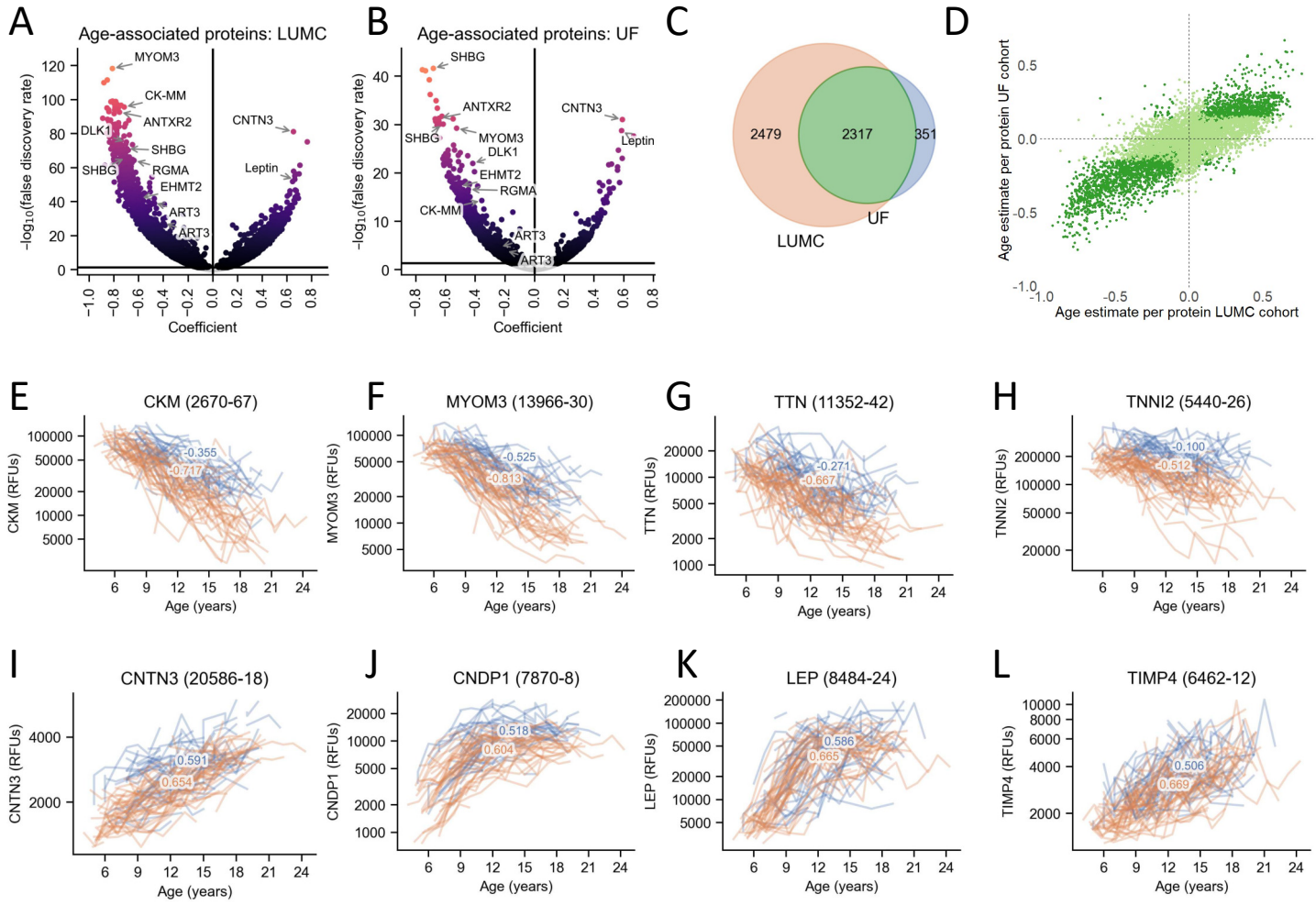
A
Age-associated proteins: LUMC
MYOM3
CK-MM
ANTXR2
DLK1
SHBG
SHBG
RGMA
EHMT2
ART3
ART3
CNTN3
Leptin
−log10(false discovery rate)
Coefficient
B
Age-associated proteins: UF
SHBG
ANTXR2
SHBG
MYOM3
DLK1
EHMT2
RGMA
CK-MM
ART3
ART3
CNTN3
Leptin
−log10(false discovery rate)
Coefficient
C
2479
2317
351
LUMC
UF
D
Age estimate per protein UF cohort
Age estimate per protein LUMC cohort
E
CKM (2670-67)
-0.355
-0.717
CKM (RFUs)
Age (years)
F
MYOM3 (13966-30)
-0.525
-0.813
MYOM3 (RFUs)
Age (years)
G
TTN (11352-42)
-0.271
-0.667
TTN (RFUs)
Age (years)
H
TNNI2 (5440-26)
-0.100
-0.512
TNNI2 (RFUs)
Age (years)
I
CNTN3 (20586-18)
0.591
0.654
CNTN3 (RFUs)
Age (years)
J
CNDP1 (7870-8)
0.518
0.604
CNDP1 (RFUs)
Age (years)
K
LEP (8484-24)
0.586
0.665
LEP (RFUs)
Age (years)
L
TIMP4 (6462-12)
0.506
0.669
TIMP4 (RFUs)
Age (years)

Figure 1

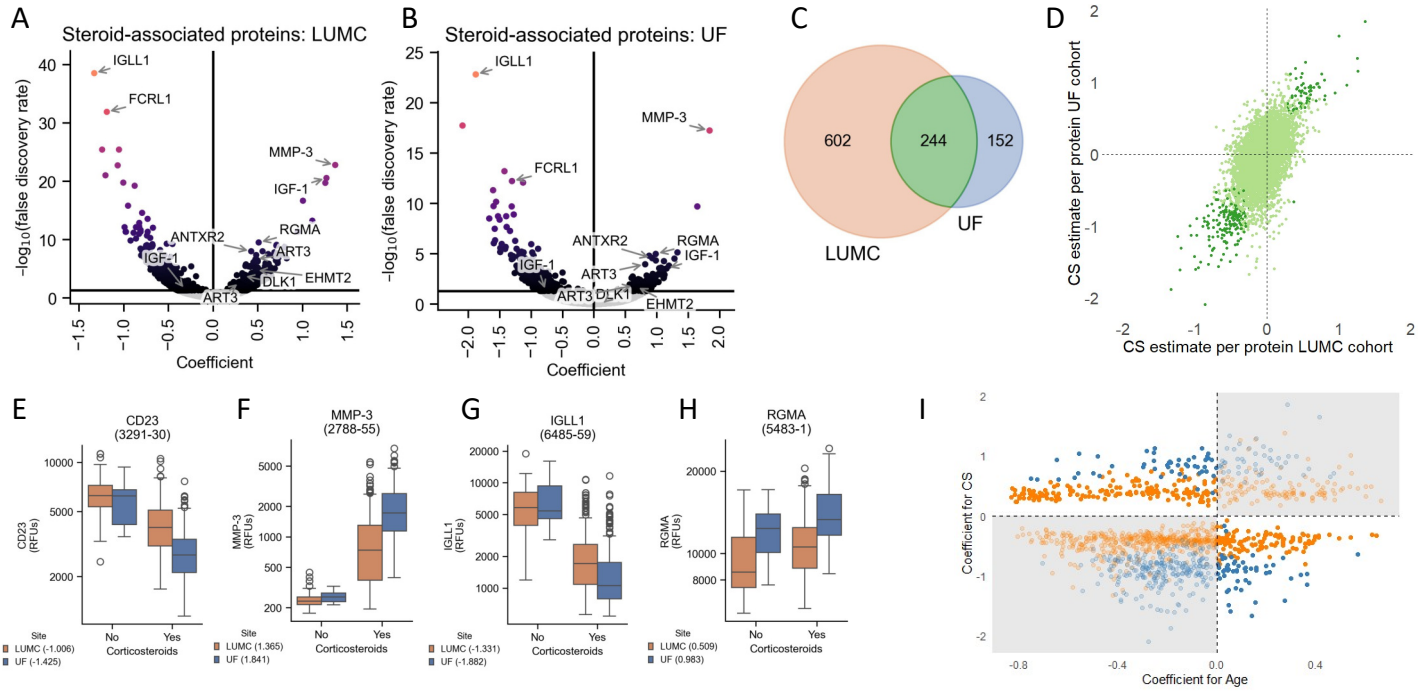
A
Steroid-associated proteins: LUMC
IGLL1
FCRL1
MMP-3
IGF-1
RGMA
ANTXR2
ART3
IGF-1
DLK1
EHMT2
ART3
−log10(false discovery rate)
Coefficient
B
Steroid-associated proteins: UF
IGLL1
MMP-3
FCRL1
ANTXR2
RGMA
IGF-1
IGF-1
ART3
ART3
DLK1
EHMT2
−log10(false discovery rate)
Coefficient
C
602
244
152
LUMC
UF
D
CS estimate per protein UF cohort
CS estimate per protein LUMC cohort
E
CD23
(3291-30)
CD23
(RFUs)
Site
LUMC (-1.006)
UF (-1.425)
No
Yes
Corticosteroids
F
MMP-3
(2788-55)
MMP-3
(RFUs)
Site
LUMC (1.365)
UF (1.841)
No
Yes
Corticosteroids
G
IGLL1
(6485-59)
IGLL1
(RFUs)
Site
LUMC (-1.331)
UF (-1.882)
No
Yes
Corticosteroids
H
RGMA
(5483-1)
RGMA
(RFUs)
Site
LUMC (0.509)
UF (0.983)
No
Yes
Corticosteroids
I
Coefficient for CS
Coefficient for Age

Figure 2

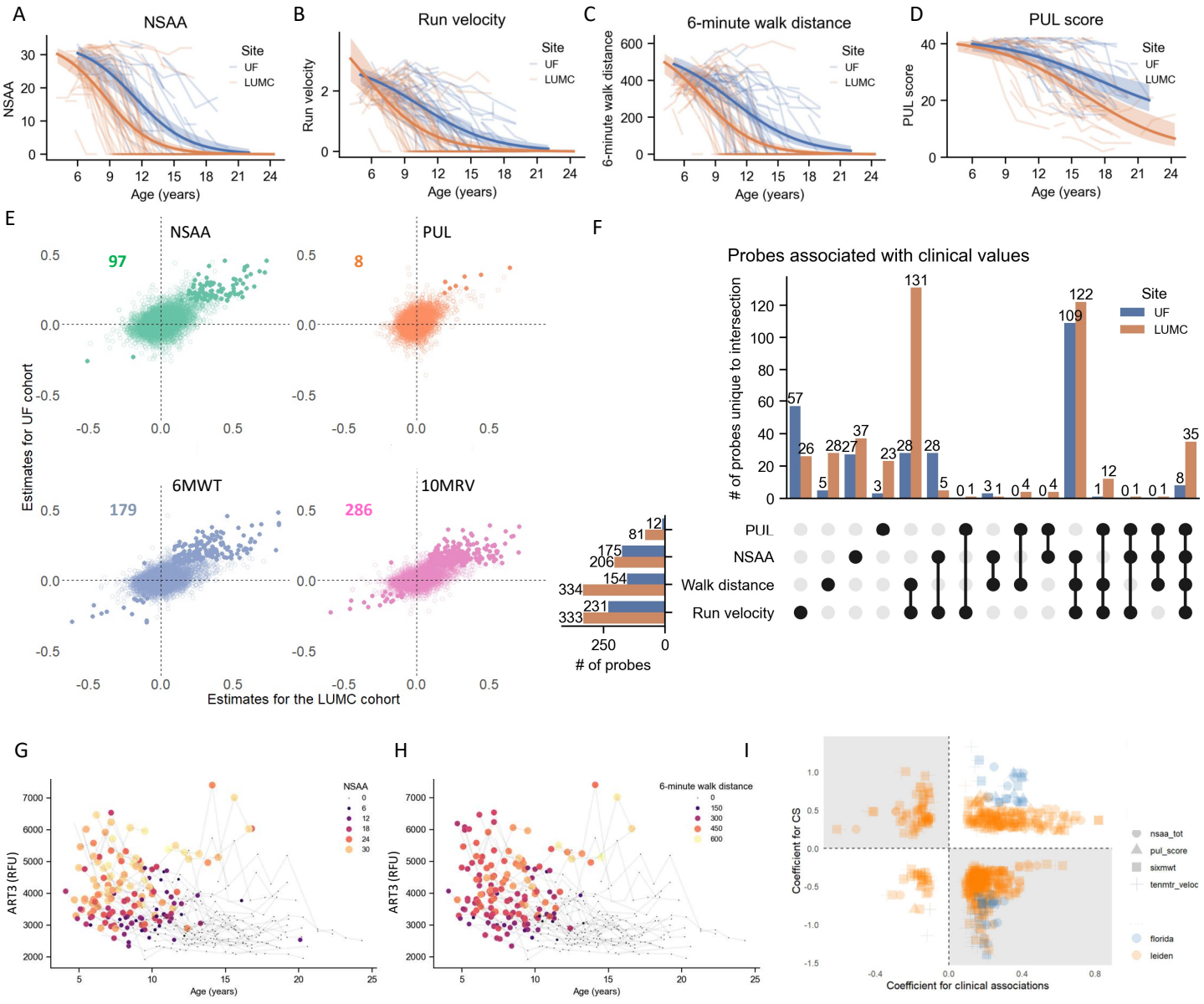
A
NSAA
B
Run velocity
C
6-minute walk distance
D
PUL score
Site
UF
LUMC
Age (years)
E
NSAA
PUL
6MWT
10MRV
97
8
179
286
Estimates for UF cohort
Estimates for the LUMC cohort
F
Probes associated with clinical values
# of probes unique to intersection
# of probes
PUL
NSAA
Walk distance
Run velocity
G
H
I
ART3 (RFU)
Coefficient for CS
Coefficient for clinical associations
nsaa_tot
pul_score
sixmwt
tenmtr_veloc
florida
leiden

Figure 3

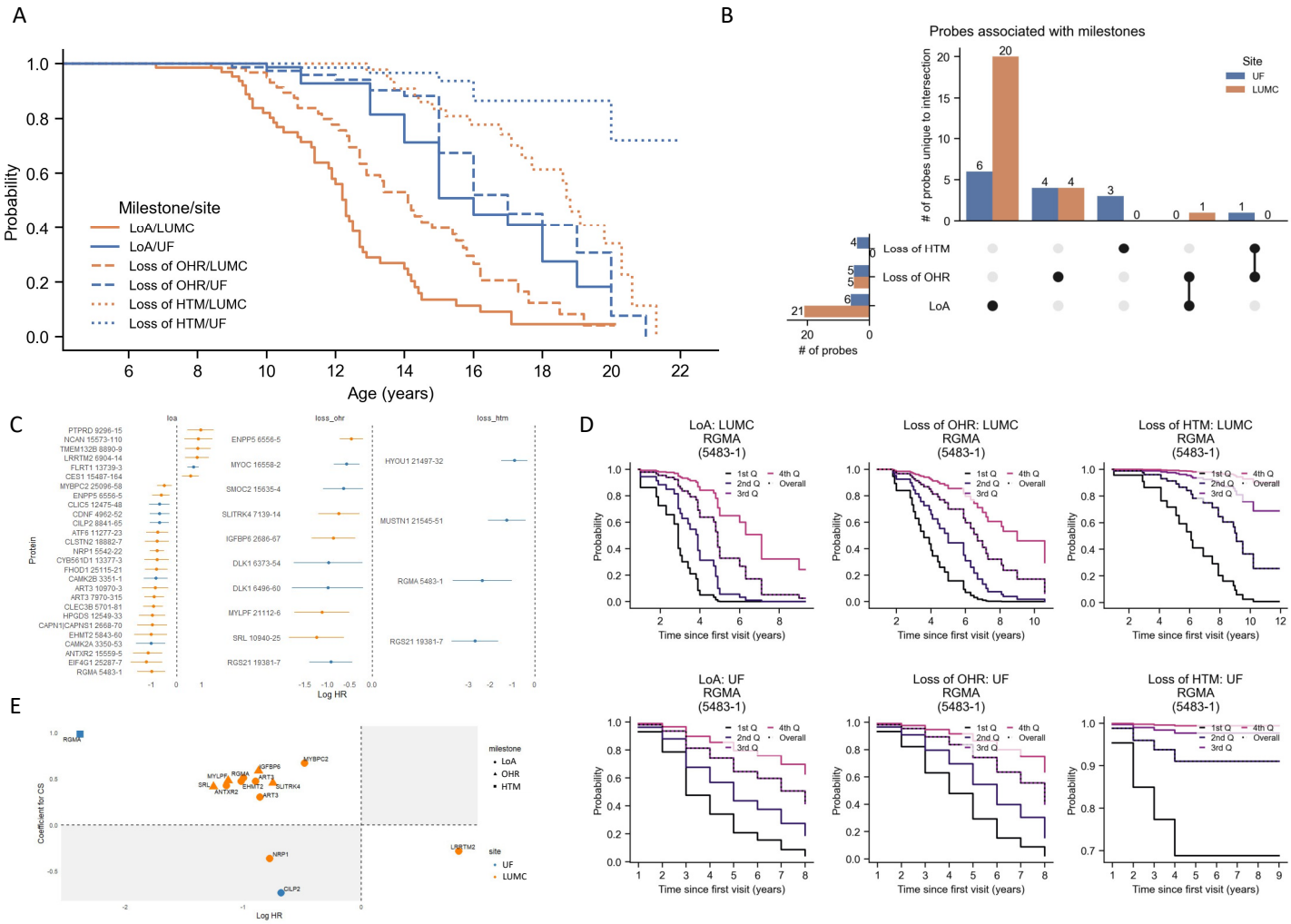
A
Probability
Age (years)
Milestone/site
LoA/LUMC
LoA/UF
Loss of OHR/LUMC
Loss of OHR/UF
Loss of HTM/LUMC
Loss of HTM/UF
B
Probes associated with milestones
# of probes unique to intersection
Site
UF
LUMC
Loss of HTM
Loss of OHR
LoA
# of probes
C
loa
loss_ohr
loss_htm
Protein
Log HR
D
LoA: LUMC
RGMA
(5483-1)
Loss of OHR: LUMC
Loss of HTM: LUMC
LoA: UF
Loss of OHR: UF
Loss of HTM: UF
1st Q
2nd Q
3rd Q
4th Q
Overall
Probability
Time since first visit (years)
E
Coefficient for CS
Log HR
milestone
LoA
OHR
HTM
site
UF
LUMC

Figure 4