






Characterisation of phenotypic patterns in equine exercise-associated myopathies

Victoria Lindsay-McGee¹  | Claire Massey¹  | Ying Ting Li¹ | Emily L. Clark²  |
Androniki Psifidi^{1,2}  | Richard J. Piercy¹ 

¹Department of Clinical Sciences and Services, Royal Veterinary College, London, UK

²The Roslin Institute, University of Edinburgh, Edinburgh, UK

Correspondence

Richard J. Piercy, Department of Clinical Sciences and Services, Royal Veterinary College, London, UK.

Email: rpiercy@rvc.ac.uk

Present address

Victoria Lindsay-McGee, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, UK.

Abstract

Background: Equine exercise-associated myopathies are prevalent, clinically heterogeneous, generally idiopathic disorders characterised by episodes of myofibre damage that occur in association with exercise. Episodes are intermittent and vary within and between affected horses and across breeds. The aetiopathogenesis is often unclear; there might be multiple causes. Poor phenotypic characterisation hinders genetic and other disease analyses.

Objectives: The aim of this study was to characterise phenotypic patterns across exercise-associated myopathies in horses.

Study design: Historical cross-sectional study, with subsequent masked case-control validation study.

Methods: Historical clinical and histological features from muscle samples ($n = 109$) were used for k -means clustering and validated using principal components analysis and hierarchical clustering. For further validation, a blinded histological study (69 horses) was conducted comparing two phenotypic groups with selected controls and horses with histopathological features characterised by myofibrillar disruption.

Results: We identified two distinct broad phenotypes: a non-classic exercise-associated myopathy syndrome (EAMS) subtype was associated with practitioner-described signs of apparent muscle pain ($p < 0.001$), reluctance to move (10.85, $p = 0.001$), abnormal gait ($p < 0.001$), ataxia ($p = 0.001$) and paresis ($p = 0.001$); while a non-specific classic RER subtype was not uniquely associated with any particular variables. No histological differences were identified between subtypes in the validation study, and no identifying histopathological features for other equine myopathies identified in either subtype.

Main limitations: Lack of an independent validation population; small sample size of smaller identified subtypes; lack of positive control myofibrillar myopathy cases; case descriptions derived from multiple independent and unblinded practitioners.

Conclusions: This is the first study using computational clustering methods to identify phenotypic patterns in equine exercise-associated myopathies, and suggests that differences in patterns of presenting clinical signs support multiple disease subtypes,

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Equine Veterinary Journal* published by John Wiley & Sons Ltd on behalf of EVJ Ltd.

with EAMS a novel subtype not previously described. Routine muscle histopathology was not helpful in sub-categorising the phenotypes in our population.

KEYWORDS

disease subtypes, EAMS, exercise-associated myopathy syndrome, exertional rhabdomyolysis, myopathy, RER

1 | INTRODUCTION

Given that horses are commonly used for athletic pursuits, it is perhaps unsurprising that exercise-associated muscle disorders are commonly recognised in this species. Often, affected horses are presumptively diagnosed with exertional rhabdomyolysis (ER), a clinically heterogeneous myopathic syndrome characterised by exercise-induced episodes of myofibre damage. Clinical signs of ER can include stiffness, muscle fasciculation, cramping, sweating, reluctance to move, myoglobinuria, and in severe cases recumbency, or death. Typically, the clinical presentations and severities of ER vary between both individuals and between episodes. ER is a rare disease in humans, where highly active people such as military personnel^{1–4} and athletes^{5–9} are commonly affected. However, ER is much more common in equivalently athletic horses: approximately 5%–7% of racing Thoroughbreds are affected.^{10–12} While some horses can have a spontaneous single episode of ER, many horses have recurring episodes: a syndrome known as recurrent ER (RER). Together with the welfare implications, this syndrome also has a substantial financial impact. One study on racehorses revealed that six training days were lost per episode, and 68% of horses were prevented from racing during the previous year.¹¹ However, beyond the racing industry, a variety of different breeds and a range of other disciplines such as harness racing,¹³ polo,¹⁴ eventing¹⁵ and endurance riding¹⁶ are also affected by exercise-associated myopathies. RER is a heritable condition in both Thoroughbreds ($h^2 = 0.34–0.46$) and Standardbreds ($h^2 = 0.39–0.49$),¹⁷ but genome-wide association studies have, so far, not found a putative causal genetic variant associated with RER.^{18,19} Similarities in disease presentation has led to assumptions that the aetiology is similar or identical between and within breeds; however, without definitive testing, this remains an assumption. Differences in presentation might be overlooked in individual cases—differences that otherwise might help define distinct subtypes.

RER syndrome is non-specific and largely idiopathic—the syndrome is generally characterised by repeated exercise-associated ER episodes of muscle stiffness or pain, and variable and intermittent post-episode elevations in muscle-derived enzyme activities. RER is also histologically non-specific, associated with evidence of myofibre damage and regeneration typical of many myopathies (i.e., presence of internalised myonuclei or myofibre size variation), but no RER-specific associated histological features. Diagnosis is often made based on exercise-association of clinical signs and on exclusion of other exercise-related myopathies such as type 1 polysaccharide storage myopathy (PSSM). Previous attempts to classify²⁰ equine RER

have instead focused on the apparent triggering factors of ER episodes. In contrast, human ER is associated with a range of distinct genetic diseases, including glycogen metabolism disorders, long-chain fatty acid metabolism disorders, mitochondrial disorders, calcium influx disorders, caveolinopathies, limb girdle muscular dystrophies, sarcoglycanopathies, dystrophinopathies and sickle cell trait.²¹ Human ER is subsequently often treated more as a symptom than a specific disease in its own right. It is possible that RER syndrome in horses could also represent a range of distinct exercise-associated myopathies or disease subtypes, particularly in different (especially unrelated) breeds.

There is precedent for identifying novel myopathies within RER syndrome in horses. In the early 1990s, Valberg et al.²² identified a subpopulation of Quarter Horses (QH) with RER that displayed a distinct histological phenotype: periodic Acid-Schiff (PAS)-positive inclusions in type 2 muscle fibres, identified as an amylase-resistant abnormal polysaccharide (polyglucosan). Various other differences in the clinical and histological phenotype of these horses with this newly identified PSSM when compared with horses with other forms of ER were identified, and additional breeds were reported with the same disease based on histopathological criteria.^{23–31} Then, in 2008 a p.R309H mutation in the glycogen synthase (GYS1) gene was identified³² and the gain of function disease mechanism subsequently explained.³³

Not all horses with a PSSM histological phenotype are positive for the p.R309H mutation.^{26,32,34,35} Consequently, those with the mutation are now designated as PSSM Type 1 (PSSM1), while those without, are designated as PSSM Type 2 (PSSM2). Currently, there is no genetic mutation nor causal disease pathway identified for PSSM2, nor for other ER-associated diseases, and the possibility remains that either, or both, PSSM2 and RER represent a syndrome comprising multiple diseases, or there is phenotypic heterogeneity (either clinical or histological) leading to phenotypically distinct disease subtypes with differing underlying aetiopathogenesis. More recently, groups of horses of certain breeds (Warmblood³⁶ and Arabian³⁷) with ER have been reported with a histopathologically distinct myopathy, termed myofibrillar myopathy (MFM), a broad term that in humans encompasses a heterogeneous group of inherited disorders with a variety of genetic causes.^{38–41} These horses have been described as having aggregates of the normal intermediate filament muscle protein, desmin, in occasional myofibres.^{36,37} It remains unclear whether these animals have a separate disease or diseases or whether they contribute to the large and apparent, ill-defined pool of affected animals. Further, and unfortunately, the confusion surrounding defining

RER phenotypes has been exacerbated by uncorroborated, but nonetheless, popular genetic testing for several of these disorders using unsubstantiated assays that currently lack validity.^{42,43}

The aim of this study was to test the hypothesis that there are subgroups of exercise-associated myopathies within horses that can be defined by clinical signs and by histopathological features in muscle biopsy samples. We used historical case records from the RVC's Comparative Neuromuscular Diseases Laboratory biopsy database and following an unbiased computational clustering analysis of the retrospective data to define phenotypes, we then performed a prospective, blinded, reassessment of histopathological features.

2 | MATERIALS AND METHODS

2.1 | Study design

In this article, we present two linked studies: Study 1, a clustering analysis of retrospective data from clinical samples in order to define phenotypic subtypes; and Study 2, a prospective, blinded validation study. Figure 1 is a flow chart presenting the steps taken in Study 1, leading to Study 2.

Briefly, in Study 1, exercise-associated myopathy cases were selected from the wider neuromuscular database, and split into three datasets, a primary dataset (Set 1) and two validation datasets (Set V1 and Set V2), each of a larger sample size but fewer variables than the primary dataset, in order to validate the findings in Set 1. These datasets underwent a primary clustering analysis of principal components analysis (PCA) and *k*-means clustering, then a secondary comparative method of hierarchical clustering. This allowed us to compare between the primary and validation datasets as well as between different computational clustering methods, to better support our findings. Variables significantly associated with the identified clusters were identified and compared. In Study 2, we selected horses from the identified clusters (potential subtypes), and these underwent a prospective, blinded, histological validation study.

2.2 | Study 1: Phenotype clustering using diagnostic biopsy service records

2.2.1 | Data curation and cleaning

Diagnostic reports for 1199 muscle biopsy and blood samples submitted to the RVC's Comparative Neuromuscular Diseases Laboratory (CNMDL) referral service from 2006 to 2017 inclusive were collated for this study. All original contemporaneous diagnoses were made by, and reports written, by a single person (co-author RP), skilled in interpretation of muscle biopsy histopathology. Signalment data were recorded for each animal, as was clinical history data as reported by the submitting veterinary surgeon, results of genetic testing where applicable, environmental factors, biochemistry and the histological features identified on the muscle biopsy sample (full list of variables

($n = 172$)) (Table S1). All clinical history and histological feature data were logged as binary variables (reported versus not reported in relation to the clinical history, and present versus not present in histological features) except for the histological disease stage (classified as acute, acute-chronic or chronic), histological severity, degree of nuclei internalisation, increase in fibre size variation (all rated on an ordinal scale from 0 to 5), and biochemistry data such as creatine kinase (CK) and aspartate amino transferase (AST) activities (binned into ordinal scores).

Horses with a diagnosis of 'idiopathic myopathy' and reported exercise association were then extracted from the database, and horses with a diagnosis of PSSM, mitochondrial myopathy, vacuolar myopathy or other defined myopathy were excluded, leaving a dataset of 205 that we then referred to as exercise-associated myopathy syndrome horses. All variables that contained no variation in these horses (such as clinical or histological signs associated with other neuromuscular diseases from the original 1199 horse biobank database, e.g., laryngeal paresis as a defining feature of recurrent laryngeal neuropathy) were dropped, as were variables with more than 50% missing data (this included PSSM1 genetic test results, electromyography and scintigraphy results). Signalment data were then masked for the automated clustering analysis, to avoid clustering based on patterns in signalment rather than disease phenotypes. After removing horses with missing data, 109 horses with both clinical history and histological features remained for further analysis. Finally, a total of 33 clinical history variables and 19 histological feature variables were included (full list of recorded and included variables can be found in Table S1). We termed this dataset Set 1.

The remaining variables were then standardised using *StandardScaler* from the *scikitlearn* library,⁴⁵ which uses a Z-scoring normalisation method whereby each variable is transformed so that the mean is 0 and the variance is 1. Then a matrix of Spearman's correlations between variables was produced using *corr* from the *pandas* library,⁴⁶ and plotted as a heat map using *heatmap* from the *seaborn* library.⁴⁷

2.2.2 | Principal components analysis and *k*-means clustering

A principal components analysis (PCA) was carried out on the Set 1 standardised clinical history and histological features data using PCA from the *scikitlearn* library⁴⁵ and an elbow plot of variance explained by principal component (PC) produced using *lineplot* from *seaborn*.⁴⁷ To select the number of centroids for *k*-means clustering, the sum of squared distances from *k*-means clustering analyses on the standardised dataset for the range of *k* from 1 to 30 were calculated and plotted as an elbow plot. *K*-means clustering was then carried out on the standardised data using the optimal identified value for *k*. The computer-assigned clusters, termed phenotypic subtypes, were then indicated on PCA score plots using *Pairgrid* from *seaborn*⁴⁷ and in a three-dimensional scatter plot using *matplotlib*.⁴⁸ Loadings were extracted for each variable, and variables with loadings >0.2 or <-0.2 on each of the first three principal components were included on biplots using *matplotlib*.⁴⁸

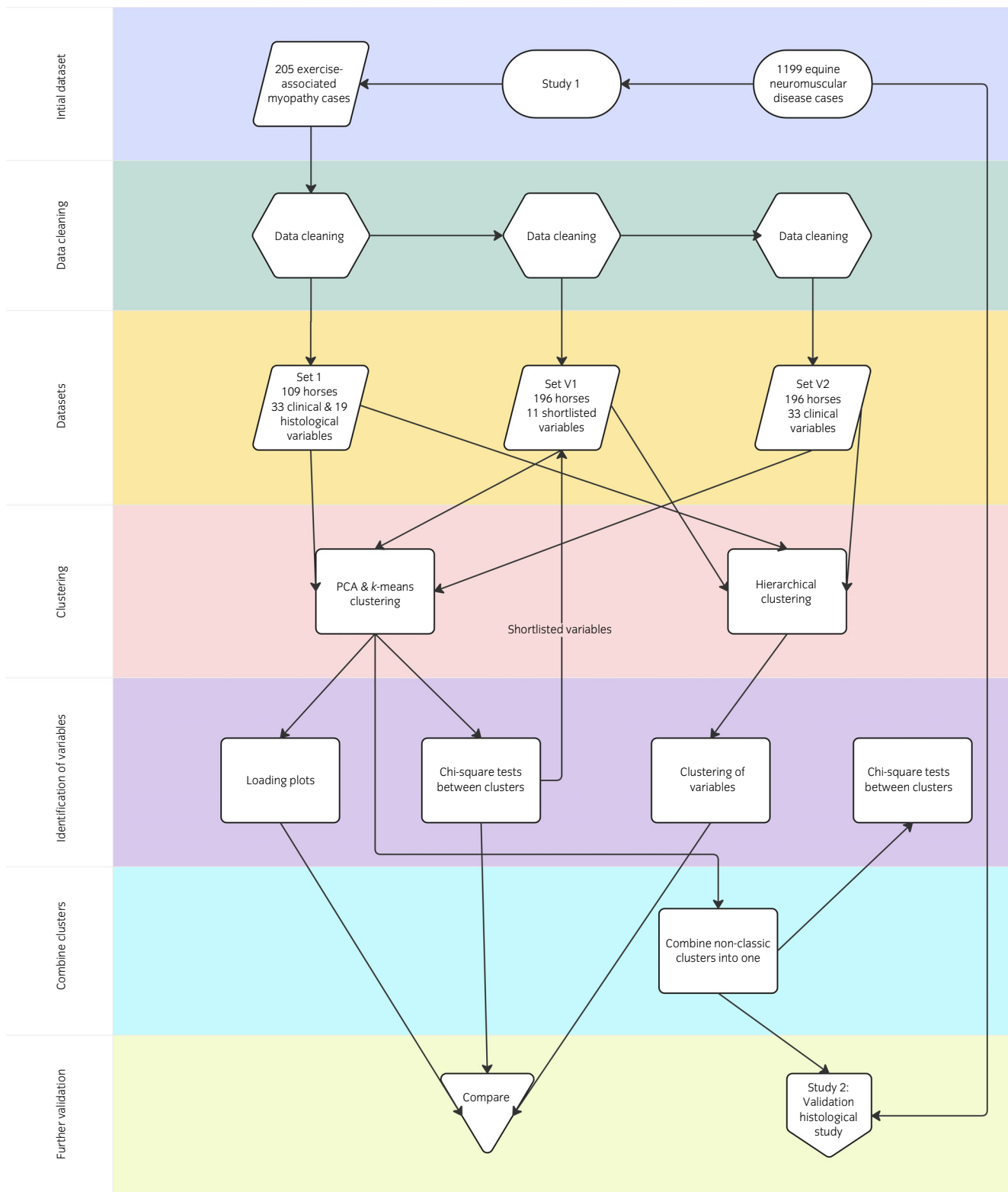


FIGURE 1 Flowchart (produced using Miro⁴⁴ software) describing the steps in Study 1, including the data curation and cleaning, the cluster analysis and identification of variables associated with distinct clusters, followed by Study 2, the validation histological study.

To identify the variables that were significantly associated with each cluster identified by the *k*-means clustering algorithm (phenotypic subtype), firstly the full list of signalment (including age, breed

and sex), clinical and histological variables was reduced from 61 to 34 using two feature selection methods to identify the best 25 variables: *SelectKBest* from scikitlearn,⁴⁵ which ranks variables based on

Chi statistics; and *ExtraTreesClassifier* from scikitlearn, which runs a random forest algorithm and produces relative feature importance scores based on the normalised total reduction in the mathematical criteria used in the decision of which feature where the tree splits. This reduced the variable list to a longlist of 34. The longlist variables were then tested individually using a Chi-squared (χ^2) test to identify important distinguishing variables between the phenotypic subtypes, with a Bonferroni-corrected significance threshold of $p < 0.0015$, and a nominal $p < 0.05$ considered suggestive (i.e., not statistically significant, but possibly worth future study in a larger cohort).

One phenotypic subtype (labelled the 'classic RER' subtype, the definition of which is discussed in more depth in the results) was then compared against the combined other groups using χ^2 tests on the same variable longlist, and the variables that were significant or suggestive, either between all four phenotypic subtypes or between classic and combined non-classic (termed exercise-associated myopathy syndrome or EAMS) subtypes, were used as the shortlisted variables for further study. Spearman's correlations were calculated between shortlisted variables and phenotypic subtypes and plotted as a heat map as above.

Differences in reported serum CK activities were also compared. Due to inconsistency in reporting from first opinion veterinary surgeons, ranging from specific CK activities at varying points post exercise or post episode, to reporting ranges or simply reporting CK activity as 'normal', and due to the unavoidable use of different laboratories and practice equipment for measuring CK activity, the data were binned into ordinal scores as detailed in Table S1. CK activity was then compared both between phenotypic subtypes (using Kruskal–Wallis test) and between classic RER and non-classic EAMS subtypes (two-sided Mann–Whitney *U*), in 49 horses.

2.2.3 | Validation using reduced input variables

To validate the robustness of the automated clustering and increase the number of horses from the initial dataset with an assigned phenotypic subtype, two larger validation datasets were also produced from the original 205 horse dataset: Set V1 ($n = 196$) consisted only of the shortlisted variables from the analysis of Set 1; Set V2 ($n = 196$) consisted of all 33 clinical history variables.

PCA and *k*-means clustering were run on both sets as described above, and significantly associated clinical signs in Set V2 were identified as described above for Set 1. Differences in serum CK activities were also compared both between phenotypic subtypes (using Kruskal–Wallis test) and between classic and non-classic EAMS subtypes (two-sided Mann–Whitney *U* test), in 98 horses.

Percentage agreement between the Set 1 clustering results (using all final histological and clinical history variables in 109 horses) and the Set V1 (using the shortlisted variables) and Set V2 (using the clinical history variables only) clustering was calculated from a contingency table, and the area under the curve (AUC) from the receiver operating characteristic (ROC) curve between the initial clustering results and the two validation sets was also compared. Spearman's

correlations between the shortlisted variables and the classic RER and non-classic EAMS subtypes from each analysis were calculated as above.

2.2.4 | Validation using hierarchical clustering

Hierarchical clustering on Set 1 plus the two validation sets (Set V1 and Set V2) was also carried out. Agglomerative clustering using Ward's method⁴⁹ of cluster similarity calculation was applied to the three sets and to the three lists of variables (all histological and clinical history variables, the shortlisted variables, and all clinical history variables, respectively) to assess grouping of variables. The results were then compared against the respective *k*-means clustering results. This analysis was carried out using *clustermap* from the seaborn library.

2.3 | Study 2: Histopathological comparison of phenotypic-based exercise-associated myopathy subtypes

As the initial reports on submitted referral biopsied muscle sample were written over a 12-year period and were performed with the observer not blinded to clinical and signalment data, and furthermore, the subjective process of assessment of biopsy samples might have altered over that period, histopathological assessments were blindly repeated and compared between horses within the classic and non-classic disease subtypes. In addition, due to similarity between some identified significant clinical signs in some of our horses and those Valberg et al. described as associated with equine MFM,^{36,37} a comparison of desmin immunohistochemistry was performed (sarcolemmal desmin aggregates being the reported defining histopathological feature for MFM^{36,37}).

2.3.1 | Sample selection

Of the 196 horses from Set V1 assigned a phenotypic subtype, we selected Thoroughbred, Warmblood or Arabian horses (Arabians and Warmbloods described as breeds associated with equine MFM^{36,37}) for which semimembranosus muscle biopsy samples were available from the CNMDL referral neuromuscular diagnostic service biobank with owner consent for research use. Due to the nature of the biobanked samples as originating from a diagnostic service for animals with clinical abnormalities, there was a lack of availability of healthy control tissue. Instead, controls were semimembranosus muscle biopsy samples from the original 1199 horse biobank dataset, where horses had no history of exercise-associated myopathy and unrelated diagnoses based on histology and clinical history: in the majority, these samples had been considered histologically normal. In addition, horses for which the original report described histological features of myofibrillar separation and/or disruption or of myofibrillar aggregates with submitted semimembranosus tissue were combined into a

'myofibrillar abnormalities' group. While desmin staining is not carried out routinely as part of the CNMDL diagnostic service and was not included for the original reports for these horses, these cases were identified as horses that otherwise might meet the description of equine MFM and could possibly act as a further positive control comparison group.

Sixty-nine horses in total were included. A description of all samples included by breed and group are summarised in Table 1.

2.3.2 | Histology

Fresh equine muscle biopsy samples submitted by practitioners from the semimembranosus muscle were frozen in isopentane, precooled in liquid nitrogen on cork discs, within 24 h, using standard methods.⁵⁰ They were kept in long term storage at -80°C . Samples were warmed to -22°C in a cryostat (Bright) before sectioning. Other than for PAS (16 μm), all sections were 8 μm in thickness. Using routine methods,⁵⁰ a panel of nine stains was applied to each sample: haematoxylin and eosin (H&E); PAS; amylase pre-digested PAS; oil red O (ORO); modified Gomori trichrome (TRI); cytochrome oxidase activity (COX); succinate dehydrogenase activity (SDH); nicotinamide adenine dinucleotide tetrazolium reductase activity (NADH); and desmin immunohistochemistry (DES IHC). The latter was conducted routinely, using an anti-desmin monoclonal mouse antibody, clone D33, Agilent (M0760), diluted 1:50 in phosphate buffered saline (PBS) incubated for 60 min. Biotinylated sheep anti-mouse IgG secondary antibody, Cytiva (RPN1001), was applied at 1:200 dilution in PBS for 45 min followed by peroxidase-conjugated streptavidin, Stratech (O16-030-084-JIR), diluted 1:500 in PBS for 30 min. Finally 3,3'-diaminobenzidine (DAB)⁵⁰ Enhanced Liquid Substrate System tetrahydrochloride, Sigma (D3939), diluted according to manufacturer's instructions was applied for 5 min.

2.3.3 | Scoring and data analysis

Histological sections were all scored by the same specialist in neuromuscular histopathology, who was blinded to group and horse signalment. The full list of assessed variables is given in Table S3. Briefly, a range of standard muscle histopathological features was ordinally scored for severity, and for distribution (whether diffuse, regional, or focal in distribution throughout the sample). Some free text variables were included, and these were analysed for themes

and assigned to dummy variables based on theme. These scores were then cleaned and analysed.

Every variable was subsequently compared across the four groups and between only the two exercise-associated myopathy groups, either using contingency tables (for categorical variables) or two-way ANOVA (for ordinal variables) with breed also compared. Variables with a p -value of less than 0.1 were retained and used in a logistic regression model (with exercise-associated myopathy group as outcome) or multinomial regression model (with disease group as the outcome) and reduced using backward stepwise elimination. This was carried out both when keeping breed dummy variables in the model, and without. These analyses were then repeated while excluding TBs from the dataset, as equine MFM has only been described in WBs and Arabians.^{36,37} Two significance thresholds were used: $p < 0.05$ for statistical significance, and $p < 0.1$ for suggestive variables possibly worth future study but not significant in this analysis.

For PCA, variables with more than 40% missing data were removed, and then all cases with missing data were dropped, leaving 47 cases and 54 variables for further analysis. The 54 variables were normalised using Z-score normalisation as detailed above, prior to conducting PCA. Loadings were then extracted for each variable, and variables with loadings >0.2 or <-0.2 on each of the first three principal components were included on biplots using matplotlib.

3 | RESULTS

3.1 | Phenotype clustering using diagnostic biopsy service records

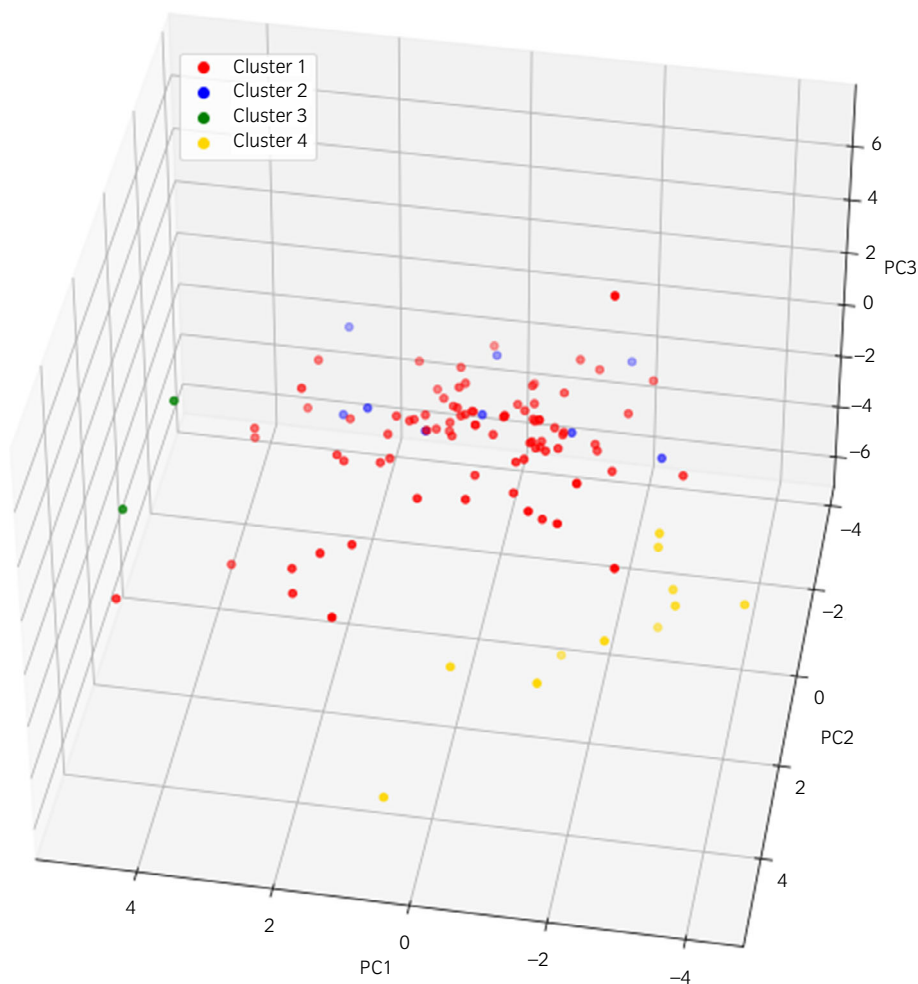
3.1.1 | Principal components analysis and k -means clustering

There was no clearly defined elbow in the elbow plot (Figure S1); however, the first point at which the slope of the trend line noticeably decreased, occurred between four and five clusters, so four clusters were selected going forward. Each horse was assigned to a phenotypic subtype using k -means clustering, and this was highlighted on a three-dimensional plot of the first three principal components (Figure 2). Phenotypic subtype 1 was the largest and most central phenotypic subtype, with PC1 separating out phenotypic subtype 3 and to a lesser extent phenotypic subtype 4, PC2 separating out phenotypic subtypes 2 and 4, and PC3 separating out phenotypic subtype 4.

TABLE 1 Samples used for Study 2: Histological comparison of exercise-associated myopathy subtypes.

Breed	Classic RER	Non-classic EAMS	Negative control	Myofibrillar defects	Total
Arabian	4	2	1	3	10
Warmblood	13	6	13	2	34
Thoroughbred	7	4	12	2	25
Total	24	12	26	7	69

FIGURE 2 Three-dimensional scatter plot of the first three principal components (PCs) from the PCA of histological and clinical history data in Set 1 ($n = 109$). Datapoints are coloured by the k -means assigned cluster (phenotypic subtype) assigned to each animal (Clusters 1–4). The first three PCs explained 6.9%, 6.1% and 5.1% of variation in Set 1, respectively. Cluster 1 represented the classic RER phenotype, and Clusters 2–4 the non-classic EAMS group.



Two feature selection methods were used, resulting in a longlist of 34 variables across groups. 64% of variables from *SelectKBest* were also identified using *ExtraTreesClassifier*. Five variables were significant using χ^2 testing at a $p < 0.05$ Bonferroni corrected threshold ($p < 0.0015$), and a further five were suggestive using a nominal $p < 0.05$ (Table S3). Phenotypic subtype 1 ($n = 87$) was not associated with any variables, but each of the other three phenotypic subtypes did have associated variables. Phenotypic subtype 2 ($n = 9$) was associated with myofibre inclusions and the neurological sign colloquially referred to as ‘shivers’^{51,52}; phenotypic subtype 3 ($n = 2$) was associated with the histological signs of whorled fibres and myofibrillar separation and/or disruption; and phenotypic subtype 4 ($n = 11$) was associated with gait abnormalities, weakness, ataxia or ataxia-like signs, reluctance to go forward and muscle pain. The Spearman’s correlations between the significant variables and the four phenotypic subtypes are shown in Figure 3.

As the lack of specific associated variables in phenotypic subtype 1 aligned with previous descriptions of RER as being non-specific in both clinical and histological signs beyond simply repeated episodes of ER, this phenotypic subtype was termed the ‘classic RER’ subtype. The other three phenotypic subtypes were

then grouped as a combined ‘non-classic’ group, termed exercise-associated myopathy syndrome (EAMS), and comparisons were repeated between the classic RER and combined non-classic EAMS subtypes (Table S3). While some histological features were significantly different between the four phenotypic subtypes, when the classic RER subtype was compared against the non-classic EAMS subtypes (consisting of phenotypic subtypes 2, 3 and 4), only clinical signs remained significant or suggestive.

Histological and clinical history variables were then compared with the loadings from the PCA, which are illustrated in biplots in Figure S3. Muscle pain, abnormal gait, stiffness, weakness and ataxia are highly influential on the first three PCs, with at least one of these variables in each PC with loadings of greater than 0.2 or less than -0.2 . Therefore, these variables account for large proportions of variation in the dataset.

No significant differences in serum CK activity scores were found ($p = 0.926$ and $p = 0.638$, respectively; Figure S2), either between all four phenotypic subtypes (Kruskal–Wallis test) and then between classic RER and non-classic EAMS subtypes (two-sided Mann–Whitney U test) in 49 horses that had suitable measurements.

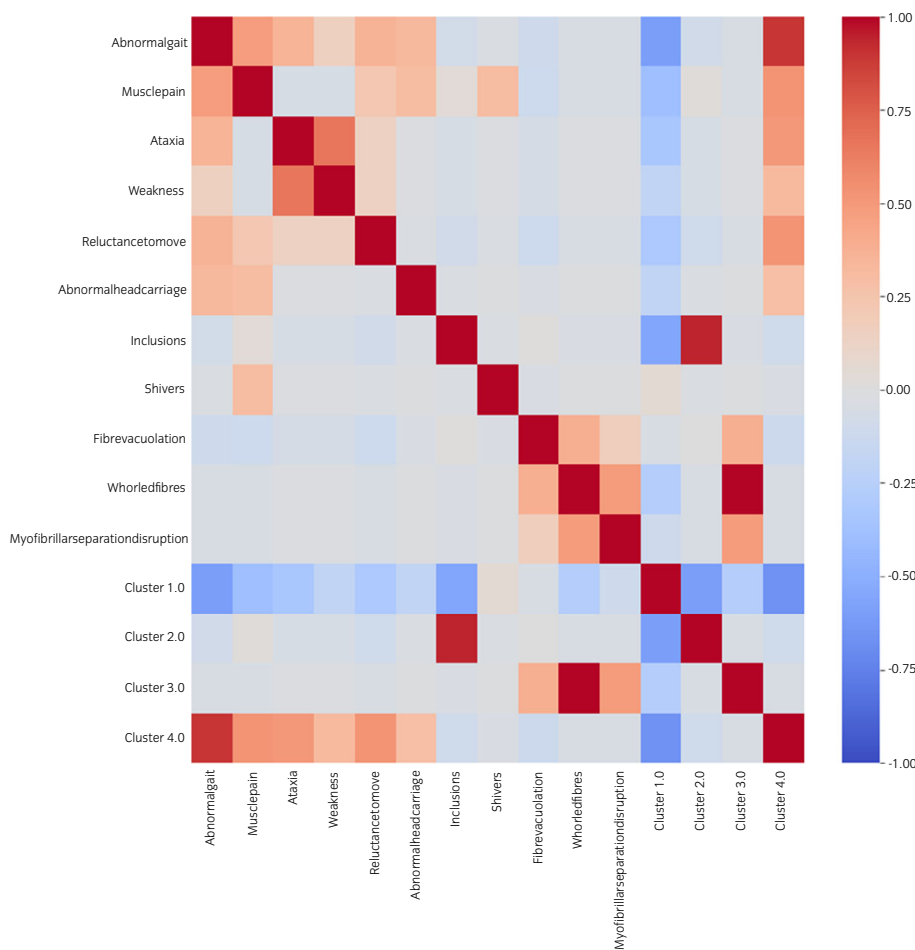


FIGURE 3 Heatmap of Spearman's correlations between the shortlisted significant and suggestive variables, and specific *k*-means assigned phenotypic subtype. Stronger positive correlations are increasingly red with stronger negative correlations increasingly blue. Cluster 1 represented the classic RER phenotype, and Clusters 2–4 the non-classic EAMS group.

3.1.2 | Validation using reduced input variables

While only 109 horses with complete signalment, clinical history and histological data were available, there were 196 horses with clinical history data available. Therefore, to validate the phenotype clustering results in the initial, well-phenotyped 109 horses, the same analyses were re-run using two validation sets based on clinical data only (Set V1, with the shortlisted clinical variables identified above, and Set V2, with all clinical variables) of these 196 horses.

In Set V1, PC1, PC2 and PC3 accounted for 16.1%, 13.7% and 13.1% of the variation in the dataset, while in Set V2, the first three PCs accounted for 5.98%, 5.91% and 5.48% of the variation, respectively. As in the Set 1 analysis, neither optimal *k* elbow plot had a decisive elbow (Figure S4). For Set V1, we therefore used the same number of phenotypic subtypes as in the Set 1 analysis (4), but in Set V2, there was a slight flattening of the trendline between 5 and 6, so 5 was used as the value for *k*. As in the Set 1 analysis, in both Set V1 and Set V2, a phenotypic subtype that had no associated variables except for ER episodes was present: in both Set V1 and Set V2 this was phenotype cluster 2, with the non-classic EAMS subtype consisting of the combined other phenotypic subtypes in each Set (biplots are shown in Figure S5).

In Set V2, variables were longlisted as for the Set 1 analysis, this time selecting the top 15 scores from each method and creating a

shortlist of 11 significant variables (Table S4). This list partially overlapped with the Set 1 associated variables, with abnormal gait, muscle pain, ataxia and shivers appearing on both lists. Presence of abnormal gait, muscle pain, ataxia, reluctance to go forward and weakness were all significantly different both between four phenotypic subtypes and between classic RER versus non-classic EAMS subtypes in the Set V1 analysis (Table S4), with whorled fibres the only significant histological feature.

The Spearman's correlations of significant variables identified from Chi-squared tests and the phenotypic subtypes identified from *k*-means clustering in both sets are presented in Figure S6. Muscle pain, abnormal gait and ataxia were also significant in these validation analyses, as were some related clinical signs in Set V2 such as poor performance and lethargy. To compare correlations across the Set 1 analysis and the two validation sets, a heatmap of Spearman's correlations of the significant clinical signs across analyses were compared against the classic RER versus non-classic EAMS subtype from each analysis (Figure S7). Broadly, the classic RER and non-classic EAMS subtypes correlated between analyses, with similar patterns in associated clinical signs.

Serum CK scores were then compared between all four phenotypic subtypes (Kruskal–Wallis test) and then between classic RER and non-classic EAMS subtypes (two-sided Mann–Whitney *U* test) in 98 and 99 horses in respective sets. These comparisons were not

significantly different in Set V1 ($p = 0.513$ and $p = 0.094$, respectively), but were when comparing between classic RER and non-classic EAMS subtypes in Set V2 ($p = 0.221$ and $p = 0.017$, respectively), with the classic RER horses having a higher CK (Figure S8).

To directly compare horses that were assigned to the classic RER or non-classic EAMS subtypes in the Set 1 analysis versus the two validation sets, contingency tables were produced (Table 2). Each validation set broadly agreed with the initial analysis but Set V1 had a higher percentage agreement at 88.1%–76.1%. Receiver operating characteristic (ROC) curves were also produced, treating the Set 1 analysis as the ‘true’ values. Area under the curve (AUC) was also slightly higher in Set V1 than Set V2 (0.79–0.78).

3.1.3 | Validation using hierarchical clustering

The hierarchical clustering of Set 1 horses using all clinical and histopathological data did not distinguish well between classic RER and non-classic EAMS subtype horses, although the first branch of the dendrogram contained notably more EAMS horses than the latter branches. However, when combined histopathological and clinical history variables were analysed, a cluster formed consisting of muscle pain, abnormal gait, reluctance to go forward and lethargy, which were identified as shortlisted variables in the k -means clustering analyses (Figure S9).

Both Set V1 and Set V2 were better distinguished between classic RER and non-classic EAMS subtypes than in the initial set using all clinical and histological data, and broad agreement was seen between k -means and hierarchical clustering (Figure S10). Similar clustering of muscle pain, abnormal gait and reluctance to go forward was noted, indicating that these clinical signs repeatedly group together across multiple datasets and clustering methods.

3.2 | Histological comparison of phenotype subgroups

Of the 69 muscle biopsy samples assessed in this study, no samples had patterns of desmin accumulation that appeared as described by

Valberg et al.^{36,37} This included all of the ‘positive control’ samples that were selected based on myofibrillar pathological features, and the non-classic EAMS subtype horses with a similar clinical phenotype to that described as equine MFM. This, supported by the finding that none of the 1096 cases submitted to the CNMDL biopsy service were given a diagnosis of MFM, indicates that the disease described by Valberg et al.^{36,37} is rare, or non-existent in UK/European horses, and distinct from the therefore novel disease phenotype described here.

In some samples, a diffuse residue from the desmin IHC was seen, rather than the protein aggregation as described in MFM. This had an appearance similar to artefactual staining which is common with DAB IHC. Table S5 illustrates Spearman's correlations between apparent desmin aggregate severity score and various key histopathology and scores for common causes of artefact in frozen muscle sections. None of these correlations was strong: the strongest being negative associations between desmin aggregate score and both freeze artefact (Spearman's $\rho = -0.320$) and fibre necrosis scores (Spearman's $\rho = -0.317$) ($p = 0.017$ and $p = 0.008$, respectively). Figure S11 illustrates Spearman's correlations between the various desmin variables and a range of other variables, including breed and disease type dummy variables. Only three variables showed moderate to high correlations with desmin variables—the ‘occasional accumulation of desmin’ pattern correlated both with freeze and with saline artefact, while the ‘central accumulation of desmin’ pattern correlated very highly with the ‘rod body’ score. However, both latter correlating variables only appeared in one horse, which happened to be the same animal—and is thus highly unlikely to be representative of any kind of a more widely applicable, causal association.

Variables were filtered for inclusion in the multinomial and logistic regression models, and backward feature elimination applied, with no variance inflation factor above 5 in any step. The final models can be seen in Table 3. None of the logistic regression models between classic RER and non-classic EAMS groups had good predictive ability (as evaluated by the pseudo- r^2 and model p -values), so the significant variables in these final models are not well supported, while in contrast, the multinomial models had low pseudo- r^2 (0.152–0.205) but were significant ($p = 0.003$ – 0.032). Across all breeds, there were no variables significantly different between the classic RER and non-classic EAMS groups, however when Thoroughbreds were excluded, fibre hypertrophy and peripheral accumulation of mitochondria

TABLE 2 Agreement between Set 1 and Set 2 classic RER versus non-classic EAMS subtypes and those of the initial analysis.

		Set V1		Set V2	
		EAMS	RER	EAMS	RER
Set 1	EAMS	14	8	18	4
	RER	5	82	22	65
Percentage agreement with classic RER subtype		88.1%		76.1%	
AUC from ROC curve between Set 1 and validation set		0.79		0.78	

Note: The classic RER subtype was the same as phenotypic subtype 1 in the Set 1 analysis and represented phenotypic subtype 2 in both Set V1 and Set V2. The non-classic EAMS subtype in the Set 1 analysis consisted of phenotypic subtypes 2, 3 and 4, while in Set V1 it consisted of phenotypic subtypes 1, 3 and 4, and in Set V2 it consisted of phenotypic subtypes 1, 3, 4 and 5.

Abbreviations: AUC, area under the curve; ROC, receiver operating characteristic.

TABLE 3 Model metrics and significant histological variables in regression models with disease group or myopathy subtype as the outcome variable.

Model	Breed	Pseudo r^2	Model p value	Final variables	Variable coefficient \pm SE (p -value)		
					Negative control	Non-classic EAMS	Myofibrillar abnormalities
Multinomial	Included	0.182	0.01	TB	0.277 \pm 0.625 (0.658)	0.240 \pm 0.755 (0.751)	-0.070 \pm 0.988 (0.943)
				Arabian	-0.770 \pm 1.207 (0.524)	-1.056 \pm 1.202 (0.380)	1.271 \pm 0.980 (0.195)
				Diffuse lobulated fibres	1.376 \pm 0.681 (0.043)	-1.458 \pm 0.903 (0.107)	0.327 \pm 1.003 (0.744)
				Myofibrillar aggregate score	2.205 \pm 1.095 (0.044)	0.315 \pm 1.323 (0.812)	2.102 \pm 1.167 (0.072)
				Polyglucosan score	-2.202 \pm 1.107 (0.047)	-1.008 \pm 1.048 (0.336)	-0.829 \pm 1.132 (0.464)
				Endomyseal lipid score	-0.704 \pm 0.440 (0.109)	-0.119 \pm 0.440 (0.755)	-1.788 \pm 0.687 (0.009)
Multinomial	Not included	0.151	0.003	Diffuse lobulated fibres	1.506 \pm 0.628 (0.017)	-1.273 \pm 0.866 (0.142)	0.265 \pm 0.898 (0.767)
				Myofibrillar aggregate score	2.251 \pm 1.063 (0.034)	0.430 \pm 1.289 (0.739)	1.893 \pm 1.103 (0.086)
				Polyglucosan score	-2.300 \pm 1.101 (0.037)	-0.990 \pm 1.042 (0.342)	-0.762 \pm 1.020 (0.455)
				Endomyseal lipid score	-0.725 \pm 0.408 (0.075)	-0.218 \pm 0.341 (0.523)	-1.483 \pm 0.608 (0.015)
Logistic	Included	No final significant variables in the model					
Logistic (WB & Arabian)	Not included	Infinite	Not estimable	Glycogen score	0.270 \pm 0.137 (0.049)		
				Arabian	-2.610 \pm 1.391 (0.061)	-2.719 \pm 1.429 (0.057)	-1.418 \pm 1.117 (0.204)
				Fibre hypertrophy score	1.397 \pm 0.644 (0.030)	1.373 \pm 0.645 (0.033)	1.226 \pm 0.639 (0.055)
				Myofibrillar aggregate score	3.167 \pm 1.530 (0.038)	1.111 \pm 1.314 (0.098)	1.403 \pm 1.047 (0.18)
				Polyglucosan score	-4.758 \pm 2.305 (0.039)	-2.004 \pm 1.266 (0.113)	-1.503 \pm 1.000 (0.133)
				Peripheral mitochondrial accumulation score	-1.386 \pm 0.649 (0.033)	-1.410 \pm 0.656 (0.031)	-2.109 \pm 0.976 (0.031)
Logistic (WB & Arabian)	Included	Infinite	>0.9	Arabian	-0.033 \pm 1.010 (0.974)		
				Internalised nuclei score	0.401 \pm 0.204 (0.049)		
Logistic (WB & Arabian)	Not included	Infinite	Not estimable	Internalised nuclei score	0.391 \pm 0.183 (0.032)		

Note: p values in bold represent variables significant at an alpha of 0.05; p values in italics represent suggestive variables at a threshold of 0.1. Abbreviation: SE, standard error.

became significant and breed became suggestive in this model. The significant variables are presented as bar plots in Figure S12.

The loading plots from the PCA are presented in Figure S13. Across the three PCs, groupings of variables are noted. Freeze artefact, saline artefact and occasional accumulation of desmin consistently group together, reflecting the correlation between these variables identified in Figure S13; whole-fibre variables such as generalised atrophy, fibre atrophy, fibre angular atrophy and fibre size variation also consistently grouped. Moreover, in some plots, fibre necrosis and peripheral accumulation of mitochondria, and subcellular variables such as polyglucosan, myofibrillar aggregates, myofibrillar loss and/or separation, and rimmed vacuoles, in some plots with rod bodies, also consistently grouped together. The datapoints themselves did not cluster by disease group.

4 | DISCUSSION

Increasingly, clinically heterogeneous complex diseases are being approached as syndromes with multiple disease subtypes that have their own genetic risk factors. In humans, studies have shown distinct phenotypic subtypes in type 2 diabetes,^{53–57} polycystic ovary syndrome,⁵⁸ preterm birth syndrome,⁵⁹ acute respiratory distress syndrome,^{60–62} and heart failure⁶³ among other complex diseases. This substantive and growing body of work indicates that this subtype clustering approach is essential for studying complex diseases. Consequently, the overall aim of this work was to determine whether distinct subtypes of exercise-associated myopathies can be defined based on the clinical history and histological data, using submissions from a referral neuromuscular diagnostic biobank, and methods that made no prior assumption of clinically important variables. Our algorithms identified a recurring phenotype, distinct from the classical RER phenotype, consisting of the clinical signs of ataxia, weakness, muscle pain, abnormal gait or reluctance to go forward under saddle that was present across clustering methods and datasets—a phenotype we have chosen to term ‘exercise-associated myopathy syndrome’ (EAMS), to distinguish this subtype from RER.

Characterising subgroups within complex disease syndromes has been uncommonly used in animals. In horses, a research group described use of PCA quantitatively to score equine metabolic syndrome (EMS) phenotypes in Arabian horses⁶⁴: the first two factor scores could predict accurately between healthy horses, high EMS risk and high pars pituitary intermedia dysfunction (PPID) risk. The particular benefit of this method was the small number of easily obtainable quantitative measures required for prediction, making this an appealing method of risk assessment in the clinic. Similarly, our study has also demonstrated a small number of clinical signs that allow exercise-associated myopathy subtypes to be simply delineated, although with the aim of classification rather than creation of a quantitative phenotype. This presents an interesting avenue for potential future work, perhaps including further use of quantitative clinically-relevant variables such as serum CK activity or other biomarkers.

In this study, we identified a consistent classic RER subtype with no significant associations with any histological or clinical history variables. Our algorithms identified at least three non-classic phenotypic subtypes in the different datasets, however the sample size within these different phenotypic subtypes was too small to confidently describe these as distinct disease subtypes; we therefore also compared the classic (RER) and combined non-classic (EAMS) subtypes. However, across multiple clustering methods (PCA, *k*-means and hierarchical clustering) and different horse groupings (the initial set was almost doubled in size in the validation sets), similar clinical signs repeatedly described EAMS.

Identification of similar described clinical signs across multiple methods with different mathematical basis and different groups of horses indicates that the signs of ataxia, weakness, muscle pain, abnormal gait and reluctance to go forward under saddle contribute to a distinct EAMS phenotype. Weakness is a common, non-specific feature of many human myopathies including the various forms of MFM,⁶⁵ and is also reported in horses with other genetic neuromuscular diseases such as hyperkalaemic periodic paralysis (HYPP).^{66,67} It is not however a phenotypic feature of ER subtypes in humans (such as those with ER-associated *RYR1* mutations⁶⁸): instead in human ER, muscle hypertrophy and improved athletic ability are common. In contrast, weakness was reported associated with EAMS in horses in this study and further, some horses were regarded as ataxic by the referring veterinary surgeon. While perceived ataxia in these horses could reflect true proprioceptive dysfunction, it might also have reflected instead paresis as the two signs can be difficult to define and distinguish, especially when signs are mild.⁶⁹ Lameness has been found in 34.1% of Standardbreds with RER previously,¹³ and while this could be partially attributed to clinical muscle pain from ER episodes, the clustering of gait abnormalities as part of a distinct phenotype indicates that there is likely something further underlying this clinical sign. Reluctance to go forward under saddle could be due to post-episode pain or subclinical muscle pain, but in most clinical histories it was implied to be a chronic problem, which would seem at odds with previous evidence of a conferred performance advantage in Standardbred horses with RER.¹³ Taken together, it would seem that these clinical features describe EAMS phenotype(s) distinct from classic RER with likely differing underlying pathophysiology and signs. We also consider it possible (perhaps probable) that this additional myopathic syndrome or syndromes, though most readily identified by owners, trainers or their attending clinician during or associated with exercise, actually represents chronic (perhaps often subclinical) disease that is also present at rest but is exacerbated by exercise; this is supported by the accompanying clinical signs involving gait, strength and/or movement.

Similar, linked clinical features have previously been reported as a possible myopathy subtype,^{36,37} in particular, in association with a disorder termed MFM. However, our blinded histological evaluation of classic and non-classic groups identified no horses with the histopathological feature of desmin aggregation, which is the defining feature of this reported disorder. As such, this non-classic EAMS subtype identified in our current study seems therefore to be novel.

However, equine MFM has previously been diagnosed following identification of as few as five or six myofibres containing desmin aggregates within an entire muscle biopsy transverse section.^{36,70} Assuming a muscle fibre diameter of 50 μm , and a biopsy section of $5 \times 5 \text{ mm}$, this equates to a proportion of desmin aggregated fibres in a single section of approximately 0.05% (1 in 2000). It is possible therefore that, if a horse in this study had as few affected fibres proportionally as previously described, the assessed muscle biopsy samples simply did not capture myofibres containing these aggregates or they were missed. However, the biological significance of so few affected fibres in a standard biopsy section is questionable: more than 75% of fibres are affected in many humans with MFM,^{38,39} and low levels of desmin and other cytoskeletal protein aggregation are often seen as non-specific features of non-desmin-related myopathies such as inclusion body myositis,⁷¹ central core⁷² and minicore myopathies.⁷³ Indeed, desmin accumulation is a typical feature of muscle regeneration⁷⁴ which would be a non-specific feature of most myopathies characterised by muscle damage and elevated serum CK activity. It seems unlikely then that horses with so few affected fibres have any specific desminopathy, or that so few affected fibres have sufficient cumulative effect to produce the severity of clinical signs described by the referring veterinary surgeons. Regardless, due to the lack of pathological desmin staining in these tested samples and considering the lack of any previously diagnosed MFM cases in this UK database, it would appear that the EAMS subtype identified in our study is not equine MFM, but instead, a distinct disease phenotype.

The only histological variables that were significantly different between classic RER and non-classic EAMS were fibre hypertrophy and peripheral mitochondrial accumulation, both of which were only significant when Thoroughbreds were excluded. It is possible, considering that breed was approaching significance between these groups, the former effect is caused by differences in exercise and training (which were not known for these samples) and the latter is linked to breed rather than disease processes. Muscle hypertrophy, or the addition of myofilaments in the myofibre, can be directly caused by training as well as by compensatory load on myofibres during disease,^{75,76} and data on exercise levels in samples from the biobank were unavailable. However, as neither fibre atrophy nor other non-specific myopathic signs were significantly different between groups, the difference in muscle fibre hypertrophy seen here is unlikely due to compensatory load. Arabian horses have a greater proportion of oxidative fibres than breeds such as Warmbloods and Thoroughbreds,⁷⁷ and unlike in humans where peripheral, subsarcolemmal mitochondrial accumulation—so called ‘ragged red’ fibres—are indicative of a mitochondrial myopathy, peripheral accumulation of mitochondria can be considered normal in horses.⁷⁸ While a significant difference between RER and EAMS subtypes could indicate a fibre type shift due to necrosis of specific fibre types and possible subsequent disease adaptations, or even some role of the mitochondria in the disease processes of these disease subtypes, this effect seems more likely simply to be attributed to fibre type differences between breeds.

In our study, none of the histological stains, nor their combination, from a typical neuromuscular diagnostic panel identified differences in histological phenotype between classic RER and non-classic EAMS disease groups. It is possible that further study using selected non-standard immunohistochemistry stains might identify differences between these groups; currently however, both subtypes have a non-specific myopathic histological appearance. This is the case in many human myopathies, including human MFM, where desminopathies, filaminopathies, ZASPopathies, $\alpha\beta$ -crystallinopathies, myotilinopathies and other MFMs have a heterogeneous and poorly-delineated histological phenotype,^{39,41,65} and in human ER, where a large range of mutations across more than 30 genes cause ER.^{68,79} Further genetic study of these distinct phenotypes might help sub-categorise idiopathic equine exercise-associated myopathies facilitating more directed histopathological comparisons.

Future prospective validation of both the non-classic EAMS subtype and the small sample size phenotypic subtypes identified in a larger, biologically independent, validation cohort would be of benefit in the next stage in characterising EAMS subtypes. As the RVC CNMDL referral neuromuscular diagnostic biobank is likely one of the largest in Europe, sourcing sufficient phenotyped European samples with a similar demographic might prove challenging. However, clinical history data should prove easier to obtain: indeed, this study suggests that clinical history is more useful for delineating phenotypic subtypes than histopathological criteria. However, independent, preferably blinded, examination by a single or small team of skilled clinicians, adept at distinguishing and accurately defining neurological and musculoskeletal gaits appears important as does future use of a standardised questionnaire for full case characterisation. However, confirmation that a horse has non-specific histopathological features that confirm the myopathy, would delineate affected horses from those with shared, vague and often poorly defined clinical signs (such as weakness) that might otherwise reflect unrelated involvement of other body systems. As such, muscle biopsy confirmation remains a key step in disease characterisation.

In summary, using clinical history and histopathological features from biobanked samples, we have identified a distinct exercise-associated myopathy phenotypic subtype we have termed EAMS, consisting of horses with weakness, perceived ataxia, muscle pain, gait abnormalities and reluctance to go forward under saddle. This was validated using different horse groupings, variable groupings, and clustering methods, with similar clinical signs consistently identified. This EAMS subtype is distinct from the disorder referred to as equine MFM; indeed, MFM was not detected in these European horses. Further, there were no histopathological features that distinguished between classic RER and non-classic EAMS across breeds, indicating that this novel subtype has a similar non-specific myopathic appearance using a standard muscle staining panel. While it remains unclear whether this novel subtype represents a single disease entity, and whether the practitioner-described clinical signs represent true associations, co-variate or simply, improper classifications, its recognition for both future genetic and clinical studies might improve study power

and help end the long delay in research progress with exertional myopathic syndromes of horses.

FUNDING INFORMATION

This work was funded by the Royal Veterinary College's Mellon Fund for Equine Research and by the RVC Comparative Neuromuscular Diseases Laboratory diagnostic biopsy service.

ACKNOWLEDGEMENTS

We thank the referring veterinary surgeons and owners for submitting samples, and David Goodwin for additional support with histology.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

Victoria Lindsay-McGee: Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization; writing – original draft; writing – review and editing. **Claire Massey:** Investigation; methodology; resources; writing – review and editing. **Ying Ting Li:** Investigation; methodology; resources; writing – review and editing. **Emily L. Clark:** Conceptualization; funding acquisition; project administration; supervision; writing – review and editing. **Androniki Psifidi:** Conceptualization; funding acquisition; methodology; project administration; resources; supervision; validation; writing – review and editing. **Richard J. Piercy:** Conceptualization; formal analysis; funding acquisition; investigation; methodology; project administration; resources; supervision; validation; writing – original draft; writing – review and editing.

DATA INTEGRITY STATEMENT

V. Lindsay-McGee had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

ETHICAL ANIMAL RESEARCH

All work was conducted with the approval of the Royal Veterinary College's Clinical Research Ethical Review Board (CRERB, reference 2018 1834-2) and Social Science Research Ethical Review Board (SSRERB, reference SR2018-1799).

INFORMED CONSENT

Explicit owner consent for animals' inclusion in the study was not stated.

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/evj.14128>.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request: Open sharing exemption granted by editor for this retrospective clinical report.

ORCID

Victoria Lindsay-McGee  <https://orcid.org/0000-0001-7693-5839>

Claire Massey  <https://orcid.org/0000-0003-0823-6547>

Emily L. Clark  <https://orcid.org/0000-0002-9550-7407>

Androniki Psifidi  <https://orcid.org/0000-0003-4095-1452>

Richard J. Piercy  <https://orcid.org/0000-0002-4344-6438>

REFERENCES

1. Tietjen DP, Guzzi LM. Exertional rhabdomyolysis and acute renal failure following the Army Physical Fitness Test. *Mil Med.* 1989;154:23–5.
2. Brown JA, Elliott MJ, Sray WA. Exercise-induced upper extremity rhabdomyolysis and myoglobinuria in shipboard military personnel. *Mil Med.* 1994;159:473–5.
3. Aizawa H, Morita K, Minami H, Sasaki N, Tobise K. Exertional rhabdomyolysis as a result of strenuous military training. *J Neurol Sci.* 1995;132:239–40.
4. Armed Forces Health Surveillance Branch. Update: exertional rhabdomyolysis, active component, U.S. Armed Forces, 2011. *MSMR.* 2012;19:17–9.
5. Clarkson PM. Exertional rhabdomyolysis and acute renal failure in marathon runners. *Sports Med.* 2007;37:361–3.
6. Kruse DH. “Tyson Squats” as a cause of rhabdomyolysis. *J Emerg Nurs.* 1998;2:116–7.
7. Mahakkanukrauh A, Sangchan A, Mootsikapun P. Exertional rhabdomyolysis following excessive exercise of university freshman cheer-training. *J Med Assoc Thai.* 2003;86:789–92.
8. Springer BL, Clarkson PM. Two cases of exertional rhabdomyolysis precipitated by personal trainers. *Med Sci Sports Exerc.* 2003;35:1499–502.
9. Oh JY, Laidler M, Fiala SC, Hedberg K. Acute exertional rhabdomyolysis and triceps compartment syndrome during a high school football camp. *Sports Health.* 2012;4:57–62.
10. MacLeay JM, Sorum SA, Valberg SJ, Marsh WE, Sorum MD. Epidemiologic analysis of factors influencing exertional rhabdomyolysis in thoroughbreds. *Am J Vet Res.* 1999;60:1562–6.
11. McGowan CM, Fordham T, Christley R. Incidence and risk factors for exertional rhabdomyolysis in thoroughbred racehorses in the United Kingdom. *Vet Rec.* 2002;151:623–6.
12. Upjohn M, Archer R, Christley R, McGowan C. Incidence and risk factors associated with exertional rhabdomyolysis syndrome in National Hunt racehorses in Great Britain. *Vet Rec.* 2005;156:763–6.
13. Isgren CM, Upjohn MM, Fernandez-Fuente M, Massey C, Pollott G, Verheyen KLP, et al. Epidemiology of exertional rhabdomyolysis susceptibility in standardbred horses reveals associated risk factors and underlying enhanced performance. *PLoS One.* 2010;5(7):e11594.
14. McGowan CM, Posner RE, Christley RM. Incidence of exertional rhabdomyolysis in polo horses in the USA and the United Kingdom in the 1999/2000 season. *Vet Rec.* 2002;150:535–7.
15. Singer ER, Barnes J, Saxby F, Murray JK. Injuries in the event horse: training versus competition. *Vet J.* 2008;175:76–81.
16. Wilberger MS, McKenzie EC, Payton ME, Rigas J, Valberg SJ. Prevalence of exertional rhabdomyolysis in endurance horses in the Pacific northwestern United States. *Equine Vet J.* 2015;47:165–70.
17. Norton EM, Mickelson JR, Binns MM, Blott SC, Caputo P, Isgren CM, et al. Heritability of recurrent exertional rhabdomyolysis in standardbred and thoroughbred racehorses derived from SNP genotyping data. *J Hered.* 2016;107(6):537–43.
18. Tozaki T, Hirota K, Sugita S, Ishida N, Miyake T, Oki H, et al. A genome-wide scan for tying-up syndrome in Japanese thoroughbreds. *Anim Genet.* 2010;41:80–6.

19. Fritz KL, McCue ME, Valberg SJ, Rendahl AK, Mickelson JR. Genetic mapping of recurrent exertional rhabdomyolysis in a population of North American thoroughbreds. *Anim Genet.* 2012;43:730–8. <https://doi.org/10.1111/j.1365-2052.2012.02351.x>
20. Keen J. Diagnosis and management of equine rhabdomyolysis. In *Pract.* 2011;33:68–77.
21. Scalco RS, Snoeck M, Quinlivan R, Treves S, Laforêt P, Jungbluth H, et al. Exertional rhabdomyolysis: physiological response or manifestation of an underlying myopathy? *BMJ Open Sport Exerc Med.* 2016; 2:e000151. <https://doi.org/10.1136/bmjsem-2016-000151>
22. Valberg SJ, Cardinet GH III, Carlson GP, DiMauro S. Polysaccharide storage myopathy associated with recurrent exertional rhabdomyolysis in horses. *Neuromuscul Disord.* 1992;2:351–9.
23. Valentine B, Credille KM, Lavoie JP, Fatone S, Guard C, Cummings JF, et al. Severe polysaccharide storage myopathy in Belgian and Percheron draught horses. *Equine Vet J.* 1997;29:220–5.
24. Sprayberry KA, Madigan J, LeCouteur RA, Valentine BA. Renal failure, laminitis, and colitis following severe rhabdomyolysis in a draft horse-cross with polysaccharide storage myopathy. *Can Vet J.* 1998;39: 500–3.
25. Bloom B, Valentine B, Glead R, Cable C. Postanaesthetic recumbency in a Belgian filly with polysaccharide storage myopathy. *Vet Rec.* 1999;144:73–5.
26. Valberg SJ, Mickelson JR, Gallant EM, MacLeay JM, Lentz L, de la Corte F. Exertional rhabdomyolysis in quarter horses and thoroughbreds: one syndrome, multiple aetiologies. *Equine Vet J.* 1999;31: 533–8.
27. Valentine B, De Lahunta A, Divers T, Ducharme N, Orcutt R. Clinical and pathologic findings in two draft horses with progressive muscle atrophy, neuromuscular weakness, and abnormal gait characteristic of shivers syndrome. *J Am Vet Med Assoc.* 1999;215:1621–65.
28. Valentine BA, McDonough S, Chang Y-F, Vonderchek A. Polysaccharide storage myopathy in Morgan, Arabian, and Standardbred related horses and welsh-cross ponies. *Vet Pathol.* 2000;37:193–6.
29. Valentine BA, Habecker PL, Patterson JS, Njaa BL, Shapiro J, Holshuh HJ, et al. Incidence of polysaccharide storage myopathy in draft horse-related breeds: a necropsy study of 37 horses and a mule. *J Vet Diagn Invest.* 2001;13:63–8.
30. Valentine B, Cooper B. Incidence of polysaccharide storage myopathy: necropsy study of 225 horses. *Vet Pathol.* 2005;42:823–7.
31. Valentine BA, Cooper BJ. Development of polyglucosan inclusions in skeletal muscle. *Neuromuscul Disord.* 2006;16:603–7.
32. McCue ME, Valberg SJ, Miller MB, Wade C, DiMauro S, Akman HO, et al. Glycogen synthase (GYS1) mutation causes a novel skeletal muscle glycogenesis. *Genomics.* 2008;91:458–66.
33. Maile CA, Hingst JR, Mahalingan KK, O'Reilly AO, Cleasby ME, Mickelson JR, et al. A highly prevalent equine glycogen storage disease is explained by constitutive activation of a mutant glycogen synthase. *BBA-Gen Subjects.* 2017;1861:3388–98. <https://doi.org/10.1016/j.bbagen.2016.08.021>
34. Stanley RL, McCue ME, Valberg SJ, Mickelson JR, Mayhew IG, McGowan C, et al. A glycogen synthase 1 mutation associated with equine polysaccharide storage myopathy and exertional rhabdomyolysis occurs in a variety of UK breeds. *Equine Vet J.* 2009;41: 597–601.
35. McCue ME, Armien AG, Lucio M, Mickelson JR, Valberg SJ. Comparative skeletal muscle histopathologic and ultrastructural features in two forms of polysaccharide storage myopathy in horses. *Vet Pathol.* 2009;46:1281–91. <https://doi.org/10.1354/vp.08-VP-0177-M-FL>
36. Valberg SJ, Nicholson AM, Lewis SS, Reardon RA, Finno CJ. Clinical and histopathological features of myofibrillar myopathy in warmblood horses. *Equine Vet J.* 2017;49:739–45. <https://doi.org/10.1111/evj.12702>
37. Valberg SJ, McKenzie EC, Eyrich LV, Shivers J, Barnes NE, Finno CJ. Suspected myofibrillar myopathy in Arabian horses with a history of exertional rhabdomyolysis. *Equine Vet J.* 2016;48:548–56. <https://doi.org/10.1111/evj.12493>
38. Selcen D, Engel AG. Myofibrillar myopathy caused by novel dominant negative α B-crystallin mutations. *Ann Neurol.* 2003;54:804–10.
39. Selcen D, Ohno K, Engel AG. Myofibrillar myopathy: clinical, morphological and genetic studies in 63 patients. *Brain.* 2004;127:439–51.
40. Vattemi G, Neri M, Piffer S, Vicart P, Gualandi F, Marini M, et al. Clinical, morphological and genetic studies in a cohort of 21 patients with myofibrillar myopathy. *Acta Myologica.* 2011;30:121–6.
41. Fichna JP, Maruszak A, Żekanowski C. Myofibrillar myopathy in the genomic context. *J Appl Genet.* 2018;59:431–9.
42. Valberg SJ, Finno CJ, Henry ML, Schott M, Velez-Irizarry D, Peng S, et al. Commercial genetic testing for type 2 polysaccharide storage myopathy and myofibrillar myopathy does not correspond to a histopathological diagnosis. *Equine Vet J.* 2021;53:690–700.
43. Valberg SJ, Henry ML, Herrick KL, Velez-Irizarry D, Finno CJ, Petersen JL. Absence of myofibrillar myopathy in quarter horses with a histopathological diagnosis of type 2 polysaccharide storage myopathy and lack of association with commercial genetic tests. *Equine Vet J.* 2022;55(2):230–8.
44. Miro. 2021. [cited 2024 Mar 20]. Available from: <https://www.miro.com>
45. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res.* 2011;12:2825–30.
46. McKinney W. Proceedings of the 9th Python in Science Conference (Austin, TX).
47. Waskom ML. seaborn: statistical data visualization. *J Open Source Softw.* 2021;6:3021. <https://doi.org/10.21105/joss.03021>
48. Hunter JD. Matplotlib: a 2D graphics environment. *Comput Sci Eng.* 2007;9:90–5.
49. Ward JH. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc.* 1963;58:236–44.
50. Dubowitz V, Sewry CA, Oldfors A. Muscle biopsy: a practical approach. Amsterdam: Elsevier Health Sciences; 2021.
51. Draper A, Bender JB, Firshman AM, Baird JD, Reed S, Mayhew IG, et al. Epidemiology of shivering (shivers) in horses. *Equine Vet J.* 2015;47:182–7.
52. Valberg SJ, Williams ZJ, Henry ML, Finno CJ. The equine movement disorder “Shivers” is associated with selective cerebellar purkinje cell axonal degeneration. *Vet Pathol.* 2015;52:1087–98.
53. Tuomi T, Santoro N, Caprio S, Cai M, Weng J. The many faces of diabetes: a disease with increasing heterogeneity. *Lancet.* 2014;383:1084–94.
54. Li L, Cheng W-Y, Glicksberg BS, Gottesman O, Tamler R, Chen R, et al. Identification of type 2 diabetes subgroups through topological analysis of patient similarity. *Sci Transl Med.* 2015;7(311):311ra174.
55. Skyler JS, Bakris GL, Bonifacio E, Darsow T, Eckel RH, Groop L, et al. Differentiation of diabetes by pathophysiology, natural history, and prognosis. *Diabetes.* 2017;66:241–55.
56. Ahlqvist E, Storm P, Käräjämäki A, Martinell M, Dorkhan M, Carlsson A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol.* 2018;6:361–9.
57. Udler MS, Kim J, von Grotthuss M, Bonàs-Guarch S, Cole JB, Chiou J, et al. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: a soft clustering analysis. *PLoS Med.* 2018;15:e1002654.
58. Dapas M, Lin FTJ, Nadkarni GN, Sisk R, Legro RS, Urbanek M, et al. Distinct subtypes of polycystic ovary syndrome with novel genetic associations: an unsupervised, phenotypic clustering analysis. *PLoS Med.* 2020;17:e1003132.
59. Barros FC, Papagorghiou AT, Victora CG, Noble JA, Pang R, Iams J, et al. The distribution of clinical phenotypes of preterm birth syndrome: implications for prevention. *JAMA Pediatr.* 2015;169: 220–9.

60. Bos L, Schouten LR, van Vught LA, Wiewel MA, Ong DSY, Cremer O, et al. Identification and validation of distinct biological phenotypes in patients with acute respiratory distress syndrome by cluster analysis. *Thorax*. 2017;72:876–83.
61. Calfee CS, Delucchi KL, Sinha P, Matthay MA, Hackett J, Shankar-Hari M, et al. Acute respiratory distress syndrome subphenotypes and differential response to simvastatin: secondary analysis of a randomised controlled trial. *Lancet Respir Med*. 2018;6:691–8.
62. Millar JE, Wildi K, Barnikowski N, Bouquet M, Hyslop K, Passmore MR, et al. Characterizing preclinical sub-phenotypic models of acute respiratory distress syndrome: an experimental ovine study. *Physiol Rep*. 2021;9:e15048.
63. Ahmad T, Lund LH, Rao P, Ghosh R, Warier P, Vaccaro B, et al. Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients. *J Am Heart Assoc*. 2018;7:e008081.
64. Lewis SL, Holl HM, Long MT, Mallicote MF, Brooks SA. Use of principle component analysis to quantitatively score the equine metabolic syndrome phenotype in an Arabian horse population. *PLoS One*. 2018;13:e0200583.
65. Selcen D. Myofibrillar myopathies. *Neuromuscul Disord*. 2011;21:161–71. <https://doi.org/10.1016/j.nmd.2010.12.007>
66. Rudolph JA, Spier SJ, Byrns G, Rojas CV, Bernoco D, Hoffman EP. Periodic paralysis in quarter horses: a sodium channel mutation disseminated by selective breeding. *Nat Genet*. 1992;2:144–7.
67. Spier S, Carlson G, Holliday T, Cardinet G 3rd, Pickar J. Hyperkalemic periodic paralysis in horses. *J Am Vet Med Assoc*. 1990;197:1009–17.
68. Scalco RS, Gardiner AR, Pitceathly RDS, Zanoteli E, Becker J, Holton JL, et al. Rhabdomyolysis: a genetic perspective. *Orphanet J Rare Dis*. 2015;10:1–15.
69. Olsen E, Dunkel B, Barker WHJ, Finding EJT, Perkins JD, Witte TH, et al. Rater agreement on gait assessment during neurologic examination of horses. *J Vet Intern Med*. 2014;28:630–8.
70. Williams ZJ, Velez-Irizarry D, Gardner K, Valberg SJ. Integrated proteomic and transcriptomic profiling identifies aberrant gene and protein expression in the sarcomere, mitochondrial complex I, and the extracellular matrix in warmblood horses with myofibrillar myopathy. *BMC Genomics*. 2021;22:438. <https://doi.org/10.1186/s12864-021-07758-0>
71. Güttches AK, Brady S, Krause K, Maerkens A, Uszkoreit J, Eisenacher M, et al. Proteomics of rimmed vacuoles define new risk allele in inclusion body myositis. *Ann Neurol*. 2017;81:227–39.
72. De Bleecker JL, Ertl BB, Engel AG. Patterns of abnormal protein expression in target formations and unstructured cores. *Neuromuscul Disord*. 1996;6:339–49.
73. Bönemann C, Thompson TG, van der Ven PFM, Goebel HH, Warlo I, Vollmers B, et al. Filamin C accumulation is a strong but nonspecific immunohistochemical marker of core formation in muscle. *J Neurol Sci*. 2003;206:71–8.
74. Bornemann A, Schmalbruch H. Desmin and vimentin in regenerating muscles. *Muscle Nerve*. 1992;15:14–20.
75. Rivero JLL. A scientific background for skeletal muscle conditioning in equine practice. *J Vet Med Ser A*. 2007;54:321–32.
76. Rivero JLL, Piercy RJ. In: Hinchcliff KW, Kaneps A, Geor R, editors. Chapter 6. Equine sports medicine and surgery. Philadelphia: Saunders; 2014. p. 69–108.
77. Snow D, Guy P. Muscle fibre type composition of a number of limb muscles in different types of horse. *Res Vet Sci*. 1980;28:137–44.
78. Valentine BA. In: Zachary JF, editor. Pathologic basis of veterinary disease. 6th ed. Missouri: Mosby; 2017. p. 908–953.e901.
79. Kruijt N, van den Bersselaar LR, Kamsteeg EJ, Verbeeck W, Snoeck MMJ, Everaerd DS, et al. The etiology of rhabdomyolysis: an interaction between genetic susceptibility and external triggers. *Eur J Neurol*. 2021;28:647–59.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lindsay-McGee V, Massey C, Li YT, Clark EL, Psifidi A, Piercy RJ. Characterisation of phenotypic patterns in equine exercise-associated myopathies. *Equine Vet J*. 2024. <https://doi.org/10.1111/evj.14128>