Original Article

Use of Cluster Analysis to Delineate Symptom Profiles in an Ehlers-Danlos Syndrome Patient Population

Jane R. Schubart, PhD, Eric Schaefer, MS, Alan J. Hakim, MA, Clair A. Francomano, MD, and Rebecca Bascom, MD

Department of Surgery (J.R.S.), College of Medicine, The Pennsylvania State University; Department of Public Health Sciences (E.S.), College of Medicine, The Pennsylvania State University, Hershey, Pennsylvania, USA; The Wellington Hospital (A.J.H.), Platinum Medical Centre, London, UK; Ehlers-Danlos Society Center for Clinical Care and Research (C.A.F.), Greater Baltimore Medical Center, Baltimore, Maryland; and Department of Medicine (R.B.), College of Medicine, The Pennsylvania State University, Hershey, Pennsylvania, USA

Abstract

Context. The Ehlers-Danlos Syndromes (EDSs) are a set of rare heritable disorders of connective tissue, characterized by defects in the structure and synthesis of extracellular matrix elements that lead to a myriad of problems including joint hypermobility and skin abnormalities. Because EDS affects multiple organ systems, defining clear boundaries and recognizing overlapping clinical features shared by disease phenotypes is challenging.

Objectives. The objective of this study was to seek evidence of phenotypic subgroups of patients with distinctive symptom profiles and describe these resulting subgroups.

Methods. Data were extracted from a repository assembled 2001–2013 by the National Institute on Aging Intramural Research Program. Agglomerative hierarchical clustering was used to form distinct subgroups of patients with respect to the domains of pain, physical and mental fatigue, daytime sleepiness, and nighttime sleep. Domains were selected based on literature review, clinician expertise, and guidance from patient advisors.

Results. One hundred seventy-five patients met all inclusion criteria. Three subgroups were identified. The Pain Dominant subgroup (39 patients) had the highest mean pain values, but lowest mean values of other symptoms. The High Symptom Burden subgroup (71 patients) had high mean values in all domains. The Mental Fatigue subgroup (65 patients) had a high mean value for mental fatigue and daytime sleepiness, but a lower mean value for pain.

Conclusion. The subgroups aligned with clinical observation of the heterogeneous nature of EDS, with overlapping symptoms between subtypes and a wide divergence in degree of symptoms within subtypes. This exploratory study helps characterize the various phenotypes and comorbidities of patients with EDS. J Pain Symptom Manage 2019;58:427–436.

Key Words

Ehlers-Danlos Syndrome, symptom profile, cluster analysis

Introduction

The Ehlers-Danlos Syndromes (EDSs) are a set of rare heritable disorders of connective tissue (HDCT) characterized by defects in the structure and synthesis of extracellular matrix elements, including collagen and proteoglycans, and the enzymes necessary to process them. Defective connective tissue can lead to a myriad of clinical problems including joint hypermobility, fragility or laxity of vasculature and visceral organs, and skin abnormalities such as hyperextensibility, velvety texture, and translucency. Clinical manifestations of EDS vary by subtype and may involve virtually every organ system in the body. Depending on the type, manifestations may include frequent joint dislocations, early-onset osteoarthritis, extensive

Address correspondence to: Jane R. Schubart, PhD, MS, MBA, Department of Surgery, College of Medicine, The Pennsylvania State University, 500 University Drive, H151, Hershey, PA 17033-0850, USA. E-mail: jschubart@pennstatehealth.psu.edu

Accepted for publication: May 24, 2019.

0885-3924/$ - see front matter
https://doi.org/10.1016/j.jpainsymman.2019.05.013

bruising, abnormal scarring, delayed wound healing, periodontal disease, arterial dissections and aneurysms, hernias, bladder and bowel dysfunction, dysautonomia, sleep disorders, fatigue, headaches, paresthesias/numbness, Chiari malformation, and scoliosis.[1]

Genetic causes have been identified for most types of EDS[1]; however, the most common type, the hypermobile type, is not understood at the molecular level, and some patients with apparent features of more rare types do not have a known genetic abnormality. The diagnostic criteria for the diverse types of EDS have evolved over the years,[1–3] and genetic testing has helped to delineate these. However, some patients do not clearly fit into specific diagnostic groups. Given this, and because EDS affects multiple organ systems, defining clear boundaries and recognizing the overlapping clinical features shared by the disease phenotypes can be challenging. Variability within disease phenotype, and even within individuals in a single family, presents an additional challenge. This challenge is complicated further by temporal variation in symptoms in a single individual over the lifespan. Robust cluster analyses of large cohorts are needed to better characterize the various phenotypes and comorbidities of this group of closely related disorders.

The primary aim of this exploratory study was to seek evidence of phenotypic clusters (subgroups of patients) with distinctive symptom profiles in a cohort of patients with EDS who completed a standardized series of questionnaires related to pain, fatigue, and sleep. Secondary aims included characterizing how the resulting subgroups differed from each other and examining the associations of each subgroup with the EDS diagnostic subtypes.

Methods

Study Sample

Data were extracted from a repository collected under a protocol entitled *Clinical and Molecular Manifestations of Heritable Disorders of Connective Tissue*, assembled between 2001 and 2013 by the National Institute on Aging Intramural Research Program, of the National Institutes of Health.[4] The participants included affected persons with a wide range of heritable connective tissue phenotypes. HDCT diagnoses were originally made according to the criteria in use at the time of enrollment and were subsequently updated in 2015. The clinical evaluation included a comprehensive history and physical examination, standardized laboratory testing, physiologic measures, and imaging. Participants completed standardized patient-reported outcome measures centered on overall health, pain, sleep, and fatigue. The clinical data set was transferred to Penn State University College of Medicine Clinical and Translational Science Institute in 2016. One of our program goals has been to perform analyses of the complex multisymptom manifestations in the HDCT to better understand the effects of these high-morbidity conditions.

In this study, we limited the analytic data set to participants in the HDCT cohort with a diagnosis of EDS. Analysis was further restricted to patients who had a physical examination at their baseline visit. A total of 252 patients in the repository met these criteria. We restricted our analysis to a subsample of patients who had complete data for at least six of the eight domains of interest (see below). When evaluating these criteria, we observed that only 22% of younger patients (age $\leq$20) met these criteria compared to 93% of older patients (age $>$20). Given the large proportion of missing data for patients $\leq$20 years old, and owing to concerns about potential selection bias, we further restricted our cohort to include only patients $\geq$21 years of age. Thus, our final cohort includes the subsample of patients diagnosed with EDS who had a physical examination at baseline, who were $\geq$21 years of age, and who had complete data for at least six domains.

Cluster Variable Measures

Eight domains were selected based on review of the literature, clinician expertise, and guidance from the Penn State EDS patient research advisory group. These domains include the following: Pain, measured by the Wisconsin Brief Pain Inventory,[5,6] using the pain severity comprehensive score; Physical Fatigue (PF) and Mental Fatigue, measured by the Multidimensional Fatigue Inventory (MFI),[7] using the PF and Mental Fatigue subscales; Daytime Sleepiness (ESS), using the Epworth Sleepiness Scale[8,9]; Nighttime Sleep (PSQI), using the Pittsburgh Sleep Quality Index[10]; Joint Hypermobility, using the Beighton score[3]; dysautonomia, using the number of symptoms listed as moderately severe or worse from the Psychological Inventory (SCL-90)[11] from the following list: headaches, nervousness/shakiness, faintness/dizziness, trembling, heart pounding/racing, feeling weak in parts of body, heavy feeling in arms or legs, and feeling afraid will faint in public; and Respiratory, using the clinical questionnaire by Sleep Medicine Associates of Maryland,[12] using number of symptoms present from the following list: shortness of breath or wheezing, nasal congestion, cough, and chest pain.

Statistical Analysis

Agglomerative hierarchical clustering[13] was used to form distinct subgroups of patients with respect to the domains of interest. The method begins by separating all patients into singleton clusters and then continues

recursively by combining the two nearest clusters into one cluster at each step. Nearness was determined by a dissimilarity matrix, which was calculated using Euclidean distances between each set of pairwise variables. Clusters were combined using Ward's method,[14] which aims to minimize the within-cluster squared deviations from the cluster mean at each step.

Before fitting the clustering algorithm, all variables were standardized by subtracting the mean value and dividing by the standard deviation to ensure that all variables were measured on the same scale. We imputed missing values using single median imputation for variables in which the percentage missing was <5% (PF subscale, daytime sleepiness, joint hypermobility, and respiratory). For variables in which the percentage missing was >5% (pain and nighttime sleep), we imputed values using recursive partitioning (RPART) models[15] to respect the underlying associations among variables. In these RPART models, all other variables were included as predictors.

The results of the fitted algorithm were graphically displayed using a dendrogram. Determination of the final number of subgroups (clusters) was based on clinical expertise and the results of the NbClust package in R,[16,17] which calculates a set of 27 metrics that attempt to determine the optimal number of subgroups under various criteria and assumptions. A multivariate analysis of variance model was also used to verify that means were jointly differentiated among subgroups for the final cluster solution. Upon formation of the final subgroups, graphical displays were created to characterize differences within and between the subgroups.

In addition, we compared variables among the subgroups that were not included in the cluster analysis. These variables were primarily from the SF-36[18] health survey (includes subscales to measure physical functioning, emotional well-being, social functioning, role limitations due to physical health problems, role limitations due to emotional problems, energy/fatigue, and general health perceptions) and also included the number of months since EDS diagnosis (reported by patients).

Results

From the repository that contained 252 EDS patients, a sample of 175 met all inclusion criteria and were retained in the analysis. All patients in the sample were at least 21 years old (mean age = 42 years), with 77% female and 86% self-identifying as white race. Enrollment in the study continued through 2014; however, most participants (82%) enrolled between 2004 and 2008 (Table 1). Patients were categorized according to the following diagnostic subtypes: Classical (26 patients), Hypermobile (34 patients), Vascular (51 patients), or Rare and Unclassified (64 patients). The Rare and Unclassified category included patients with the rarer forms of Ehlers-Danlos Syndrome, including kyphoscoliotic and arthrochalasia subtypes, and patients with features that overlapped with two or more subtypes and were therefore difficult to classify.

With respect to the eight domains, 103 patients (59%) had no missing values, 55 patients (31%) had only one missing value, and 17 patients (10%) had two missing values. Nighttime sleep (22% missing) and pain (17% missing) were missing in a large percentage of patients. All other variables had <5% missing. Missing values were imputed as described in the Methods section and used in the cluster analysis.

Fig. 1 shows scatter plots and Pearson's correlation coefficient for each pair of variables (complete cases only for each pairwise comparison). No obvious strong associations were apparent, and correlation coefficients generally indicated small to moderate associations. Only daytime sleepiness and PF subscale (correlation = 0.52), daytime sleepiness and respiratory (correlation = 0.47), and PF subscale and pain (correlation = 0.43) had a correlation coefficient >0.40. Based on these results, joint hypermobility, dysautonomia, and respiratory were excluded from further consideration in the clustering analysis. Dysautonomia and respiratory were excluded because, unlike all other domains, the variables were not derived from validated scales but were counts of specific symptoms, as reflected in the highly skewed distributions (most patients had scores of 0 or 1). Joint hypermobility was excluded because it was skewed toward large values, as expected for patients with EDS, and tended to have small associations with other variables (correlation coefficients ranged from 0.09 to 0.23). After these exclusions, the cluster analysis focused on the five remaining domains.

Fig. 2 shows the dendrogram of the cluster solution based on the remaining five domains: pain, physical fatigue, mental fatigue, daytime sleepiness, and nighttime sleep. Based on the results of NbClust, two-cluster and three-cluster solutions were both reasonable as these were selected by eight and six metrics, respectively, of the 27 metrics calculated. We selected the three-cluster solution because the subgroups aligned with clinical experience and judgment. Furthermore, the multivariate analysis of variance results indicated that means were jointly different ($P < 0.0001$) for the three-cluster solution. We named the subgroups as follows: Pain Dominant (shown in red in the figures), which contained 39 patients (22%); High Symptom Burden (blue), which contained 71 (41%); and Mental Fatigue (green), which contained 65 (37%).

Table 1
Demographics, Domains, and Other Important Variables Stratified by Subgroups

Variables	Pain Dominant ($n = 39$)	High Symptom Burden ($n = 71$)	Mental Fatigue ($n = 65$)	Total ($N = 175$)
Demographics				
Age				
n	39	71	65	175
Mean (SD)	43.1 (12.3)	41.6 (11.6)	42.8 (12.5)	42.4 (12.0)
Gender				
Female	25 (64.1%)	61 (85.9%)	49 (75.4%)	135 (77.1%)
Male	14 (35.9%)	10 (14.1%)	16 (24.6%)	40 (22.9%)
Race				
Asian or Pacific Islander	0 (0%)	0 (0%)	1 (1.5%)	1 (0.6%)
Black or African American	1 (2.6%)	0 (0%)	0 (0%)	1 (0.6%)
Hispanic	0 (0%)	5 (7%)	1 (1.5%)	6 (3.4%)
White	35 (89.7%)	59 (83.1%)	56 (86.2%)	150 (85.7%)
Unknown/not reported	3 (7.7%)	7 (9.9%)	7 (10.8%)	17 (9.7%)
Year of enrollment				
2004 to 2008	34 (87.2%)	59 (83.1%)	50 (76.9%)	143 (81.7%)
2009 to 2014	5 (12.8%)	12 (16.9%)	15 (23.1%)	32 (18.3%)
Domains used in cluster analysis				
Pain score				
n	20	64	62	146
Mean (SD)	6.2 (1.1)	5.7 (1.4)	2.8 (1.3)	4.5 (2.0)
Physical fatigue score				
n	39	70	65	174
Mean (SD)	11.2 (4.0)	17.6 (2.4)	11.8 (4.2)	14.0 (4.6)
Mental fatigue score				
n	39	71	65	175
Mean (SD)	7.8 (3.3)	14.0 (3.3)	10.9 (3.9)	11.5 (4.3)
Nighttime sleep (PSQI score)				
n	31	54	52	137
Mean (SD)	5.6 (3.1)	11.5 (4.3)	6.7 (3.1)	8.3 (4.4)
Daytime sleepiness (ESS score)				
n	37	69	65	171
Mean (SD)	7.9 (4.6)	9.6 (5.3)	7.8 (4.4)	8.5 (4.9)
Other important variables				
Beighton score				
n	39	70	65	174
Median (IQR)	5 (3–8)	6 (5–8)	6 (3–7)	6 (4–8)
Dysautonomia (number of symptoms)				
n	38	66	62	166
Median (IQR)	1 (0–2)	2 (1–4)	1 (0–2)	1 (0–3)
Respiratory (number of symptoms)				
n	38	67	63	168
Median (IQR)	0 (0–1)	1 (0–2)	1 (0–1)	1 (0–2)
Time since diagnosis (months)				
n	26	61	49	136
Median	26 (14–48)	30 (12–58)	15 (5–32)	24 (10–48)
SF36: Physical functioning score				
n	37	71	64	172
Mean (SD)	57.8 (34.0)	34.2 (22.4)	66.8 (25.0)	51.4 (30.0)
SF36: Role limitations due to physical health score				
n	37	70	65	172
Mean (SD)	53.4 (45.3)	10.8 (24.4)	57.7 (41.4)	37.7 (42.6)
SF36: Role limitations due to emotional problems score				
n	37	70	65	172
Mean (SD)	85.6 (26.7)	55.7 (42.0)	77.9 (36.5)	70.5 (39.0)
SF36: Energy/fatigue score				
n	35	71	65	171
Mean (SD)	47.6 (27.1)	23.1 (16.2)	45.9 (22.1)	36.8 (23.9)
SF36: Emotional well-being score				
n	35	70	65	170
Mean (SD)	74.3 (14.5)	62.1 (18.8)	73.7 (17.1)	69.0 (18.2)
SF36: Social functioning score				
n	36	71	65	172
Mean (SD)	66.0 (27.7)	39.6 (27.9)	71.5 (25.1)	57.2 (30.5)
SF36: General health score				
n	36	71	65	172
Mean (SD)	56.3 (21.6)	27.5 (15.4)	50.4 (22.2)	42.2 (23.1)

SF-36 Lower scores indicate more disability. Values for each variable are shown on the original scale.

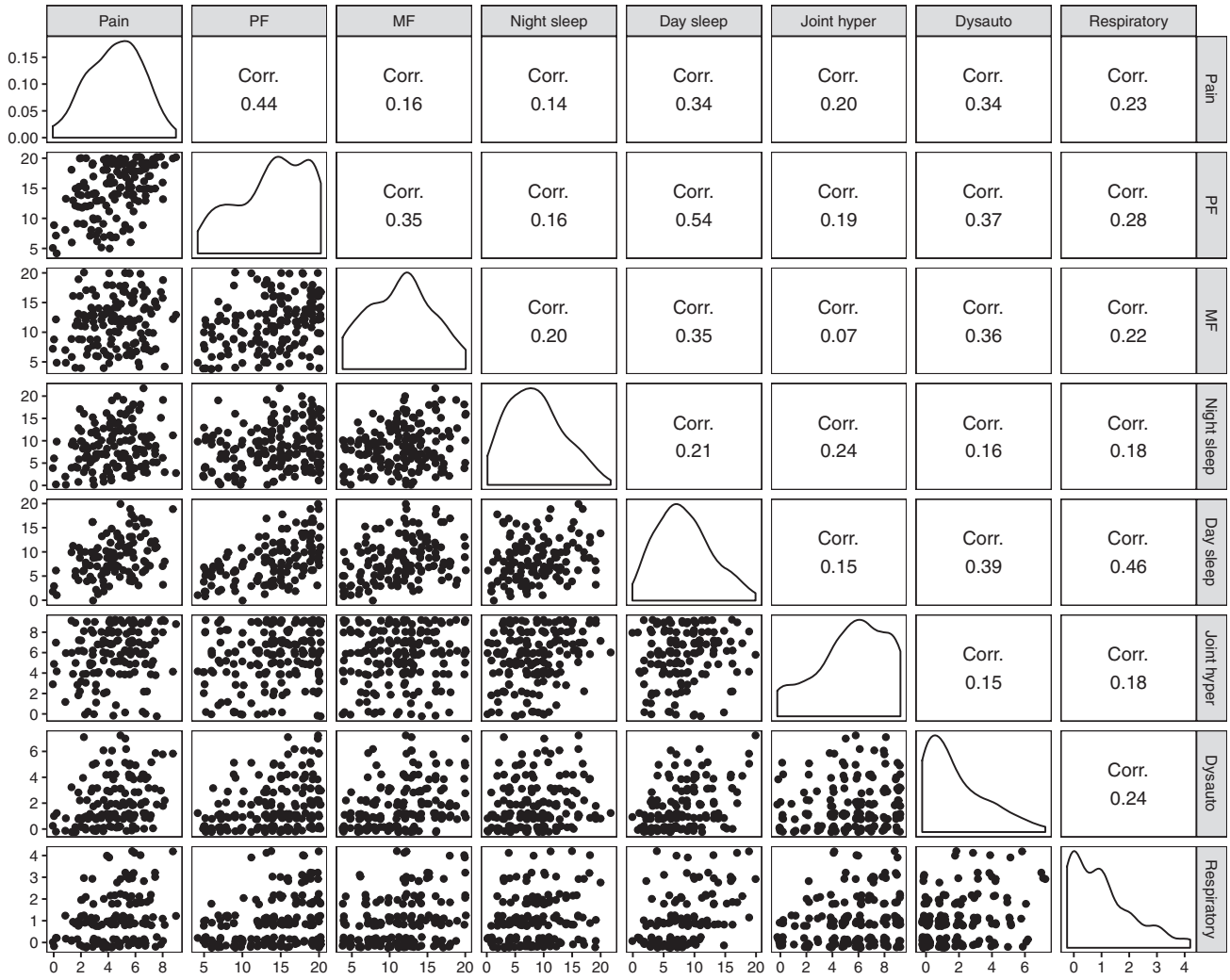

Fig. 1. Scatterplot matrix for each pair of variables, with density estimates shown on diagonal and Spearman's correlation statistics shown in upper corner. Points were slightly jittered (offset) within the scatter plots.

Fig. 3a shows the standardized means for each variable by subgroup. The Pain Dominant subgroup was characterized by the highest value of mean pain scores among the subgroups, but the lowest mean values of all other symptoms. The High Symptom Burden subgroup was characterized by high mean values in all variables. The Mental Fatigue subgroup was characterized by a higher mean value for mental fatigue and daytime sleepiness relative to the other variables but also had a very low mean value for pain relative to the other subgroups. Notably, mean values for daytime sleepiness were similar between the subgroups. Fig. 3b shows the values of the variables (standardized) for each patient within each subgroup. Each subgroup was characterized by large amounts of variation across all variables, although differences were still noticeable between subgroups. For example, the largest values of pain (on the standardized scale) for the Mental Fatigue subgroup were approximately 0.5, compared to values >2 for the other subgroups.

Composition of the Cluster Analysis Subgroups

The composition of each subgroup was also examined to determine whether patients with the same EDS diagnostic subtype clustered together. Table 2 shows the subgroups by EDS subtype diagnosis. The Pain Dominant subgroup occurred in similar proportions in each of the EDS diagnostic subtypes (between 19 and 24% of persons with each diagnosis) and, therefore, does not appear to be specific to a subtype of EDS. The High Symptom Burden subgroup was found in all diagnostic subtypes, although disproportionately higher in the Classical and Hypermobile subtypes (58% and 53%, respectively) compared to Vascular (31%) and Rare/Unclassified (34%). Conversely, Mental Fatigue was more likely to occur in Vascular (47%) and Rare/Unclassified (42%) compared to Classical (23%) and Hypermobile (24%).

Examination of additional variables among the subgroups is shown in Table 1. The Pain Dominant and

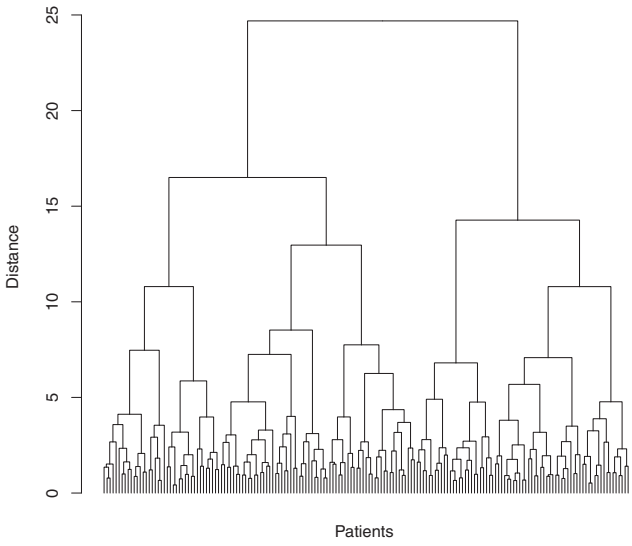

Fig. 2. Final dendrogram based on agglomerative hierarchical clustering. The leaves on the very bottom of the plot indicate individual patients. Line segments at various heights (y-axis) indicate the distances at which clusters were combined.

Mental Fatigue subgroups had similar mean Beighton values (5.1 and 5.2, respectively), compared to the High Symptom Burden subgroup (6.1). Patients in the High Symptom Burden subgroup were slightly younger but reported longer duration (in months) since learning of their EDS diagnosis and greater disability on SF-36 scores, including physical functioning, social functional, emotional functioning, energy/fatigue, role limitations due to physical health, role limitations due to emotional problems, and general health. These patients also reported more respiratory symptoms.

Discussion

The primary aim of this exploratory study was to seek evidence of phenotypic clusters (subgroups) with distinctive symptom profiles in a cohort of patients with EDS who completed a standardized series of questionnaires related to pain, fatigue, and sleep. The cluster analysis was blinded to the diagnosed EDS subtype. The study was motivated by the authors' clinical observations of heterogeneity within diagnostic subtypes and cross-cutting similarities between persons of different EDS subtypes.

The main finding is the identification of three subgroups. We find these subgroups to be recognizable, useful, and thought-provoking. They are recognizable because they align with our clinical observation that, within the multitude of symptoms and heterogeneous spectra of EDS, there are symptom profiles or "clinical themes" that emerge over time. Also recognizable is the wide interindividual variation in symptom severity and the presence of each EDS subtype within each subgroup, aligning with the observation that there is vast phenotypic variability of the EDS subtypes and clinical overlap between EDS subtypes.[1] Although exploratory, the subgroups are useful to clinicians in drawing attention to possible endophenotypes and pointing out the limitations of stereotyping within EDS diagnostic subtypes. They are thought-provoking because they draw attention to symptom patterns that may have distinct pathophysiologies. In particular, the presence of high and low pain phenotypes challenges the notion that pain is an invariable feature of EDS. Although the underlying mechanisms of pain in EDS are not well understood, both acute pain and chronic pain are common manifestations.[19]

A secondary aim was to describe differences among these empirically derived EDS subgroups. The Pain Dominant subgroup reports high levels of pain, but relatively low physical and mental fatigue and sleep disturbance. It is possible that this group has a specific intense peripheral pain source related to subluxations/dislocations.[20] Understanding the sources of their pain and resilience factors (e.g., exercise) will be important to understanding this group.

The High Symptom Burden cluster is characterized by both high pain and high fatigue (both physical and mental) and sleep disturbances (i.e., nonrestful sleep and daytime sleepiness). In clinical practice, fatigue is common and often disabling.[21–24] A 2019 study, using retrospective clinical data from chart review and patient-reported survey questionnaire data of 47 children and adolescents with hypermobile EDS and hypermobility spectrum disorders, found that general fatigue and pain scores were the best predictors of quality of life.[25] A prominent feature of some patients with EDS is multifocal pain that is not fully explained by injury or inflammation of joints.[22,26–28] In these EDS patients who experience chronic pain states, peripheral nociceptive input may be responsible for some of the patient's pain, but central nervous system factors (e.g., neurotransmitters, inflammatory pathways) amplify the pain. Because pain pathways are amplified, pain is experienced throughout the body—widespread musculoskeletal pain, chronic headaches, visceral pain. Dr. Daniel Clauw has noted that these central factors may also contribute to fatigue, memory problems (brain fog), and sleep and mood disturbances, possibly because the same neurotransmitters that control pain and sensory sensitivity also control sleep, mood, memory, and alertness.[29,30] Because pain is common in EDS and associated with functional impairment,[22,23,31] understanding centralized pain in EDS patients is important for clinicians in selecting treatments such as surgery or opioids.[32]

The Mental Fatigue subgroup is characterized by low relative pain. However, these patients report

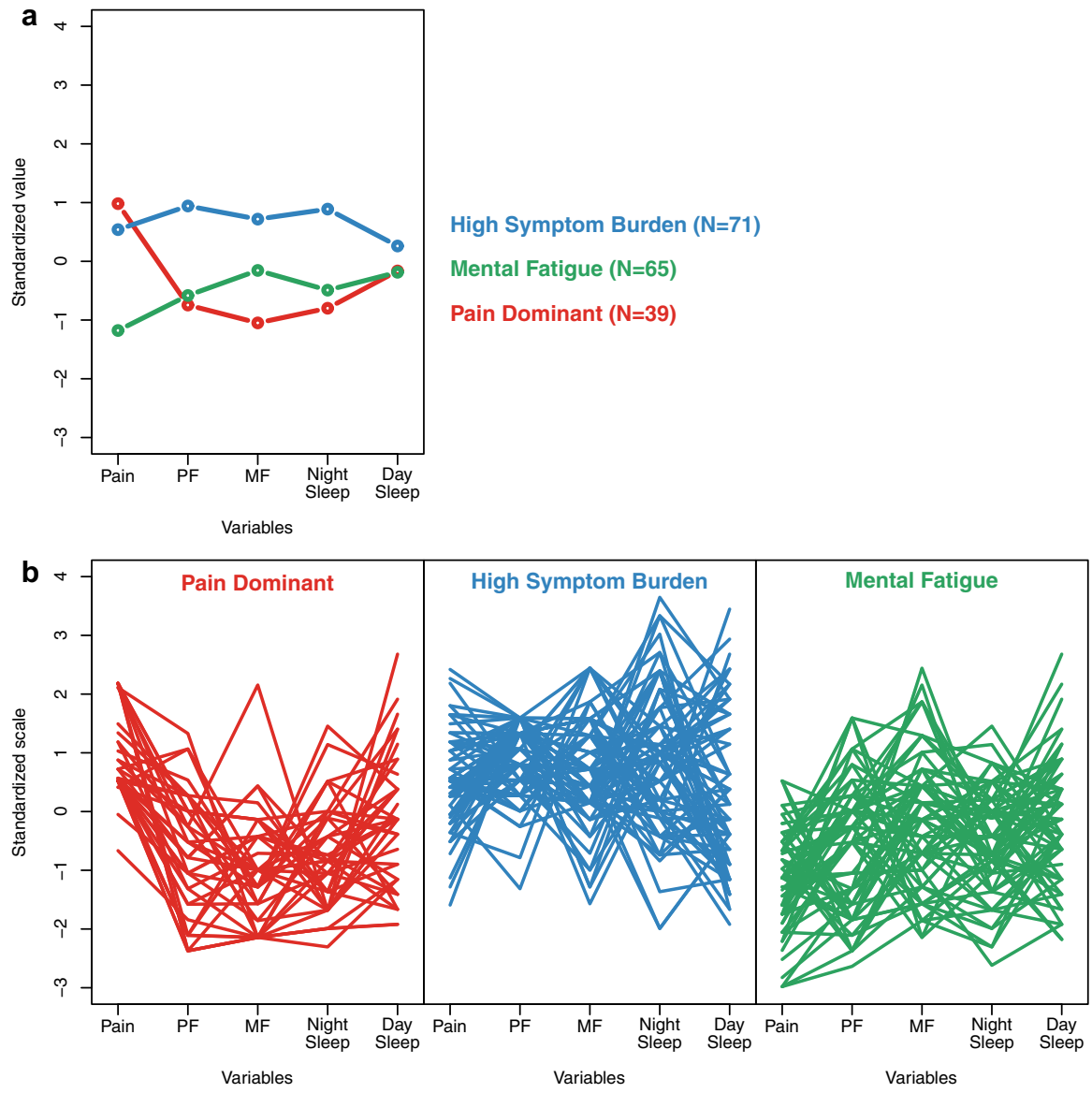

Fig. 3. a) Mean values by subgroup for each standardized variable. b) Parallel coordinates plots showing standardized values across variables for all individuals classified within each subgroup.

mental fatigue and daytime sleepiness on structured surveys that may remind clinicians of brain fog and sleepiness complaints. The sleepiness implies disturbed sleep, possibly disrupting a downregulator of the pain amplification pathway, and yet they report lower pain. Clauw identifies sleep and exercise as two things that help turn down the amplifier.[33] The questionnaires in our study did not ask about exercise, and so we do not know whether this might be a mitigating factor in this group.

We also evaluated the Beighton score among the clusters because it is the accepted standard for

Table 2
Cluster Subgroups by Diagnosed EDS Subtype

Subgroups	Classical ($n = 26$), %	Hypermobile ($n = 34$), %	Vascular ($n = 51$), %	Rare/unclassified ($n = 64$), %
1. Pain dominant (red)	5 (19.2)	8 (23.5)	11 (21.6)	15 (23.4)
2. High symptom burden (blue)	15 (57.7)	18 (52.9)	16 (31.4)	22 (34.4)
3. Mental fatigue (green)	6 (23.1)	8 (23.5)	24 (47.1)	27 (42.2)

EDS = Ehlers-Danlos Syndrome.

measuring generalized joint hypermobility, which is a common feature of EDS.[27,28] The Beighton score measures the presence of hyperlaxity at both wrists, the fifth metacarpophalangeal joints, elbows, knee joints, and the lumbosacral spine, and sums the items to yield a score ranging from 0 to 9 points.[34] The Beighton score was not highly correlated with other domains measured in this analysis, which was surprising because we presumed that the origin of pain in EDS was related to the presence of hypermobility. However, a recognized limitation of the Beighton score is that it does not include hips, shoulders, and neck and therefore does not give a complete picture of an individual's hypermobility. Alternate explanations are that the hypermobility-related pain relationship is modified by unidentified factors or that the overall level of hypermobility (i.e., the Beighton score) is not as important as the degree of hypermobility at one or several joints.

Another secondary aim was to examine the association of subgroups with EDS subtypes. Before conducting the analysis, we expected that empirically derived subgroups would be aligned with previously identified, genetically determined EDS subtypes. (There is not a genetic signature identified for hypermobile EDS.) Our results do not support these expectations. The percentage of participants in the Pain Dominant subgroup was similar in all EDS diagnostic subtypes, while the High Symptom Burden subgroup was more prevalent in the Classical and Hypermobile subtypes and the Mental Fatigue subgroup was more prevalent among the Vascular and Rare/Unclassified subtypes. Thus, pain would appear to be neither specific nor sensitive with respect to use in diagnostic criteria distinguishing between the subtypes of EDS. These results challenge the notion that persons with the hypermobile subtype of EDS are more likely to present with pain[23,35,36] and that the vascular subtype is less likely to present with this complication. Byers et al. note that muscle and tendon rupture appears to be increasing in vascular EDS and that pain also results from repetitive subluxations due to hyperlaxity of the temporomandibular articulation.[37] These data also motivate the search for environmental and genetic factors that determine the high pain in some EDS patients.

A closer examination of variables among the subgroups suggests that higher overall symptom burden and disability are associated with duration of disease. Patients in the High Symptom Burden subgroup had a similar median age to other subgroups, but the median time since learning of their diagnosis was longer. While many patients live with high symptom burden before diagnosis of EDS, it is reasonable to speculate that those with higher symptom burden would have sought treatment at a younger age. This highlights the need for longitudinal studies to examine symptom changes, aging, and meaningful clinical and quality of life outcomes.

Our study has several limitations. The features we selected as the inputs for the cluster analysis were limited to what was available within the repository. Another limitation is that statistical imputation methods were used to replace missing values with estimates (imputation), although this is preferable to analyzing only the complete cases, which might lead to biased results. Another limitation is that the hierarchical clustering algorithm, by its very nature, will force clusters to exist even in instances when the data lie along a continuum and do not exhibit distinct clustering. For this reason, the ultimate success or effectiveness of the clustering results needs to be made after evaluating the statistical results along with clinical intuition and expertise. The clinician coauthors found the subgroups generated by the cluster analysis to align with their experience. Another limitation of the study is that the participants in this study were not diagnosed according to the 2017 nosology, but rather by the Villefranche criteria that preceded the most recent diagnostic algorithms.[3]

Strengths of the study include a data-driven objective assessment of a large cohort for EDS, a rare disease, and the use of a well-characterized cohort with ascertainment using standardized instruments for the cluster analysis. However, the instruments did not measure all the features that we would like to include.

Hierarchical cluster analysis is an exploratory approach that identifies groups of objects (e.g., individuals) or variables (e.g., symptoms) based on similarity between them.[38] As such, our study should be considered exploratory in nature. Cluster analysis has been applied to identify clinical phenotypes other disciplines, including chronic obstructive pulmonary disease,[39] complex chronic pain,[39] fibromyalgia,[40] adult-onset diabetes,[41] and cancer.[42] The concept of symptom clusters in oncology has been advanced by the work of Dodd et al.[43] and Miaskowski et al.,[44] suggesting that concurrent and related symptoms may or may not be related through a common mechanism or etiology[45] (e.g., inflammatory pathways) and that symptom clusters may produce different outcomes than individual symptoms.[46] Symptom cluster analysis has been used to examine how symptoms are related to one another (e.g., fatigue, depression, sleep disturbance, and pain) and to suggest treatment strategies that could simultaneously impact clustered symptoms (e.g., acupressure, cognitive-behavioral therapy).[47] Similar to our exploratory work to delineate symptom profiles in EDS, patient subgroups with a similar symptom experience or biomarkers have also been studied in

oncology and there is some evidence that these subgroups endure over time.[46]

To our knowledge, our study is the first application of agglomerative hierarchical clustering to delineate symptom profiles and characterize phenotypic differences in a cohort that includes patients with all EDS subtypes. In patients with the hypermobile EDS subtype and hypermobility spectrum disorders, Copetti et al. (2019) identified two distinct severity groups using an approach that included hierarchical clustering to explore variables generated by principal components analysis.[48] Future studies would ideally include additional standardized instruments to look at possible mitigating factors (e.g., exercise), medication use, and features of EDS that are proposed by some as comorbid phenotypic characteristics. These include gastrointestinal dysmotility, mast cell dysfunction, and dysautonomia, among many others.[19] As Miaskowski notes in her work in symptom experiences of oncology patients, understanding the biological and behavioral mechanisms is important to develop and test interventions.[49] Another parallel to our work in EDS, as noted by Dodd et al.[50] in oncology patients, longitudinal studies are needed to validate the subgroups by examining whether disease varies over time by subgroups with respect to important outcomes such as symptoms, response to treatment, and disease progression.

Disclosures and Acknowledgments

The project described was supported by the National Institute on Aging Intramural Research Program and by the National Center for Advancing Translational Sciences, National Institutes of Health, United States, through grant UL1 TR002014. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

The authors have no conflicts of interest to disclose.

References

1. Malfait F, Francomano C, Byers P, et al. The 2017 international classification of the Ehlers-Danlos syndromes. Am J Med Genet C Semin Med Genet 2017;175:8–26.

2. Beighton P. Ehlers-Danlos syndrome. Ann Rheum Dis 1970;29:332–333.

3. Beighton P, De Paepe A, Steinmann B, Tsipouras P, Wenstrup RJ. Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-danlos National Foundation (USA) and ehlers-danlos support group (UK). Am J Med Genet 1998;77:31–37.

4. Bascom R, Schubart JR, Mills S, et al. Heritable disorders of connective tissue: description of a data repository and initial cohort characterization. Am J Med Genet A 2019; 179:552–560.

5. Cleeland CS. Measurement of pain by subjective report. In: Chapman CR, Loeser JD, eds. Advances in Pain Research and Therapy, 12. New York, NY: Raven Press, 1989:391–403.

6. Tan G, Jensen MP, Thornby JI, Shanti BF. Validation of the brief pain inventory for chronic nonmalignant pain. J Pain 2004;5:133–137.

7. Smets EM, Garssen B, Bonke B, De Haes JC. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. J Psychosom Res 1995;39: 315–325.

8. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991;14:540–545.

9. Johns MW. Reliability and factor analysis of the Epworth sleepiness scale. Sleep 1992;15:376–381.

10. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193–213.

11. Derogatis LR, Cleary PA. Confirmation of the dimensional structure of the SCL-90: a study in construct validation. J Clin Psychol 1977;33:981–989.

12. Pulmonary and Critical Care Associates of Baltimore (PCCAB). Sleep disorders. Available from https://pccab.com/services/sleep-disorders/. Accessed February 13, 2019.

13. Hastie T, Tibshirani R, Friedman J. Hierarchial clustering. In: Hastie T, Tibshirani R, Friedman J, eds. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York, NY: Springer, 2016:520–528.

14. Ward JH Jr. Hierarchical grouping to optimize an objective function. J Am Stat Assoc 1963;58:236–244.

15. Therneau T, Atkinson B. rpart: Recursive partitioning and regression trees. R package version 4.1-13. 2018. Available from https://CRAN.R-project.org/package=rpart. Accessed February 10, 2019.

16. Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: an R package for determining the relevant number of clusters in a data set. J Stat Softw 2014;61:1–36.

17. R Core Team. R. A language an environment for statistical computing. 2015. Available from https://www.r-project.org/contributors.html. Accessed April 30, 2019.

18. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473–483.

19. Tinkle BT, Malfait F, Francomano CA, Byers PH. Special issue: the ehlers-danlos syndromes: reports from the International Consortium on the Ehlers-Danlos sydromes. Am J Med Genet C Semin Med Genet 2017;175:1–245.

20. Castori M. Pain in Ehlers-Danlos syndromes: manifestations, therapeutic strategies and future perspectives. Expert Opin Orphan Drugs 2016;4:1145–1158.

21. Castori M, Camerota F, Celletti C, et al. Natural history and manifestations of the hypermobility type Ehlers-Danlos syndrome: a pilot study on 21 patients. Am J Med Genet A 2010;152A:556–564.

22. Scheper MC, de Vries JE, Verbunt J, Engelbert RH. Chronic pain in hypermobility syndrome and Ehlers-Danlos syndrome (hypermobility type): it is a challenge. J Pain Res 2015;8:591–601.

23. Voermans NC, Knoop H, Bleijenberg G, van Engelen BG. Pain in ehlers-danlos syndrome is common, severe, and associated with functional impairment. J Pain Symptom Manage 2010;40:370–378.

24. Verbraecken J, Declerck A, Van de Heyning P, De Backer W, Wouters EF. Evaluation for sleep apnea in patients with Ehlers-Danlos syndrome and Marfan: a questionnaire study. Clin Genet 2001;60:360–365.

25. Mu W, Muriello M, Clemens JL, et al. Factors affecting quality of life in children and adolescents with hypermobile Ehlers-Danlos syndrome/hypermobility spectrum disorders. Am J Med Genet A 2019;179:561–569.

26. Castori M, Morlino S, Celletti C, et al. Re-writing the natural history of pain and related symptoms in the joint hypermobility syndrome/Ehlers-Danlos syndrome, hypermobility type. Am J Med Genet A 2013;161A:2989–3004.

27. Murray B, Yashar BM, Uhlmann WR, Clauw DJ, Petty EM. Ehlers-Danlos syndrome, hypermobility type: a characterization of the patients' lived experience. Am J Med Genet A 2013;161A:2981–2988.

28. Castori M, Tinkle B, Levy H, Grahame R, Malfait F, Hakim A. A framework for the classification of joint hypermobility and related conditions. Am J Med Genet C Semin Med Genet 2017;175:148–157.

29. Clauw DJ. Fibromyalgia: a clinical review. JAMA 2014; 311:1547–1555.

30. Clauw DJ. Fibromyalgia. In: Mayer E, Bushnell M, eds. Functional Pain Syndromes: Presentation and Pathophysiology. Seattle, WA: IASP Press, 2009:580.

31. Sacheti A, Szemere J, Bernstein B, Tafas T, Schechter N, Tsipouras P. Chronic pain is a manifestation of the Ehlers-Danlos syndrome. J Pain Symptom Manage 1997;14:88–93.

32. Nijs J, Malfliet A, Ickmans K, Baert I, Meeus M. Treatment of central sensitization in patients with 'unexplained' chronic pain: an update. Expert Opin Pharmacother 2014; 15:1671–1683.

33. Clauw DJ. Chronic pain - is it all in their head?. 2013. Available from https://www.youtube.com/watch?v=pgCfkA9RLrM. Accessed April 10, 2019.

34. Naal FD, Hatzung G, Muller A, Impellizzeri F, Leunig M. Validation of a self-reported Beighton score to assess hypermobility in patients with femoroacetabular impingement. Int Orthop 2014;38:2245–2250.

35. Tinkle B, Castori M, Berglund B, et al. Hypermobile Ehlers-Danlos syndrome (a.k.a. Ehlers-Danlos syndrome Type III and Ehlers-Danlos syndrome hypermobility type): clinical description and natural history. Am J Med Genet C Semin Med Genet 2017;175:48–69.

36. Bowen JM, Sobey GJ, Burrows NP, et al. Ehlers-Danlos syndrome, classical type. Am J Med Genet C Semin Med Genet 2017;175:27–39.

37. Byers PH, Belmont J, Black J, et al. Diagnosis, natural history, and management in vascular Ehlers-Danlos syndrome. Am J Med Genet C Semin Med Genet 2017;175: 40–47.

38. Aktas A, Walsh D, Hu B. Cancer symptom clusters: an exploratory analysis of eight statistical techniques. J Pain Symptom Manage 2014;48:1254–1266.

39. Bruehl S, Harden RN, Galer BS, Saltz S, Backonja M, Stanton-Hicks M. Complex regional pain syndrome: are there distinct subtypes and sequential stages of the syndrome? Pain 2002;95:119–124.

40. Vincent A, Hoskin TL, Whipple MO, et al. OMERACT-based fibromyalgia symptom subgroups: an exploratory cluster analysis. Arthritis Res Ther 2014;16:463.

41. Ahlqvist E, Storm P, Karajamaki 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;6:361–369.

42. Miaskowski C, Cooper BA, Paul SM, et al. Subgroups of patients with cancer with different symptom experiences and quality-of-life outcomes: a cluster analysis. Oncol Nurs Forum 2006;33:E79–E89.

43. Dodd M, Janson S, Facione N, et al. Advancing the science of symptom management. J Adv Nurs 2001;33: 668–676.

44. Miaskowski C, Dodd M, Lee K. Symptom clusters: the new frontier in symptom management research. J Natl Cancer Inst Monogr 2004:17–21.

45. Kelly DL, Dickinson KA, Hsiao CP, et al. Biological basis for the clustering of symptoms. Semin Oncol Nurs 2016;32: 351–360.

46. Barsevick A. Defining the symptom cluster: how far have we come? Semin Oncol Nurs 2016;32:334–350.

47. Kwekkeboom KL. Cancer symptom cluster management. Semin Oncol Nurs 2016;32:373–382.

48. Copetti M, Morlino S, Colombi M, Grammatico P, Fontana A, Castori M. Severity classes in adults with hypermobile Ehlers-Danlos syndrome/hypermobility spectrum disorders: a pilot study of 105 Italian patients. Rheumatology (Oxford) 2019. Available from https://doi.org/10.1093/rheumatology/kez029. [Epub ahead of print], pii: kez029.

49. Miaskowski C. Future directions in symptom cluster research. Semin Oncol Nurs 2016;32:405–415.

50. Dodd MJ, Cho MH, Cooper BA, Miaskowski C. The effect of symptom clusters on functional status and quality of life in women with breast cancer. Eur J Oncol Nurs 2010; 14:101–110.