Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Investigating reliability and internal consistency of the Spider, a multisystemic symptom screening tool for people with hypermobility related disorders

Ellen Ewer

ellenewer@gmail.com

University College London

Henry W W Potts

University College London

Robby De Pauw

Ghent University

Inge De Wandele

Ghent University Hospital

Jane V. S Simmonds

University College London

Research Article

Keywords: The Spider

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11182017/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations:
No competing interests reported.

Table 4 is available in the Supplementary Files section.

Investigating reliability and internal consistency of the Spider, a multisystemic symptom screening tool for people with hypermobility related disorders.

Abstract

Background: Hypermobility Spectrum Disorders (HSD) and hypermobile Ehlers Danlos Syndrome (hEDS) are often associated with symptoms affecting many bodily systems. The Spider questionnaire assesses eight common symptoms: pain, fatigue, neuromusculoskeletal, gastrointestinal, urogenital, cardiac dysautonomia, anxiety, and depression. This study evaluated the Spider's reliability and internal consistency.

Methodology: Ethical approval was obtained. Hypermobility support organisations aided recruitment through social media advertisements. Participants self-selected based on inclusion criteria and completed two online surveys, one-to-two weeks apart. Data collection included demographics, the Spider questionnaire and the SF-36/PEDS-QL. Relative reliability was assessed using Intraclass Correlation Coefficients (ICC). Absolute reliability was evaluated using the Standard Error of Measurement (SEM), Minimal Detectable Change (MDC), and Bland-Altman analysis. Pearson correlations with SF-36/PedsQL scores were compared across timepoints to monitor stability. Internal consistency was analysed using Cronbach's alpha (α).

Results: 617 participants aged 13-65 with HSD/hEDS or symptomatic hypermobility completed both surveys. Test-retest reliability was good to excellent in individual domains and total score (ICC: 0.79 to 0.92). SEM ranged from 4.0 to 11.60 and MDC from 11.08 to 32.13. Seven domains and the total score demonstrated acceptable internal consistency (α=0.70 to 0.91). The pain domain had marginally lower internal consistency (α=0.69).

Conclusions: The Spider questionnaire demonstrated good to excellent test-retest reliability. Seven of the eight domains and the total score demonstrated acceptable internal consistency. These findings support the reliability of Spider scores for repeated measurement over time.

Key words:

The Spider

Hypermobility

Multisystemic screening tool

Reliability

Key points:

- This article presents the assessment of reliability of the Spider tool, adding to the body of work investigating its psychometric properties.
- The research shows that the Spider is a reliable tool, with acceptable test-retest reliability across all domains, and acceptable internal consistency in seven domains and the total score.
-
-
-
-
-

Background

The reported prevalence of generalised joint hypermobility (GJH) in the general population ranges between 2-57% [1]. GJH is often asymptomatic, but 10% of people report physical and psychological symptoms [1]. These symptoms most commonly include pain, joint instability and musculoskeletal injuries, however around two-thirds of people also report symptoms extending beyond the musculoskeletal system affecting multiple body systems [2]. The mechanisms for the multisystemic involvement are not yet fully understood, although several hypotheses have been proposed [2]. It is thought that an individual's presentation is influenced by the interaction of genetics, developmental and functional interactions, and the environment [3].

Some patients with symptomatic GJH are diagnosed with an underlying disorder, such as Hypermobility Spectrum Disorders (HSD), or hypermobile Ehlers Danlos Syndrome (hEDS). HSD and hEDS can be described as existing on a continuum, sharing similar symptom presentations and management approaches. These conditions are differentiated during diagnosis based on organ fragility, skin symptoms and/or skeletal morphology [4, 5]. Symptoms associated with HSD/hEDS include pain, fatigue, neuromusculoskeletal symptoms, cardiac dysautonomia, gastrointestinal and urogenital symptoms and anxiety and depression. The presence of these symptoms has been investigated and is summarised in Table 1 [6-15]. The multisystemic nature of these conditions has been shown to influence quality of life (QoL), disease severity, disability and health status, and if left unmanaged can lead to poor prognosis [16-21]. Management should be individualised following a thorough subjective and objective assessment regardless of diagnosis [2].

Table 1: Common symptoms associated with HSD/hEDS.

Symptom		Prevalence
Pain	Neuropathic pain	78% [6]
	Generalised soft tissue or body pain	90% [6]
	Chronic musculoskeletal pain	89% [7]
	Joint pain	83.2% [8]
	Headache	60.5% [8]
	Migraine	42.5% [8]
Neuromusculoskeletal	Dislocations	78% [6]
	Subluxations	65.9% [8]
Fatigue	Chronic fatigue	77% [9]
Gastrointestinal	Abdominal pain	69% [10]
	Postprandial fullness	34% [10]
	Constipation	73% [10], 63.8% [8]
	Diarrhoea	47% [10], 62.3% [8]
	Nausea	76.6% [8]
Cardiac Dysautonomia	POTS	60-80% [11]
	Orthostatic hypotension, POTS and uncategorised orthostatic intolerance	78% [12]
Urogenital	Urinary incontinence	68-84% [13]
	Genital pain and discomfort	60% [13]
Mental Health	Anxiety	70% [14], OR 4.4 [14]
	Depression	OR 4.1 [15]

Key: *POTS* – postural orthostatic tachycardia syndrome, *OR* – odds ratio

The symptom profiles of individuals with HSD/hEDS are often complex and varied and can change over time [22], making individualised care and monitoring of multisystemic symptom changes paramount. A self-report questionnaire, 'The Spider', has been developed and validated by a multidisciplinary team comprising clinicians and researchers who specialise in HSD/hEDS. The Spider's 31 questions concisely assess the presence and impact of eight common comorbid conditions associated with HSD/hEDS, as described above. The development of the Spider has been described in previous work by Ewer and colleagues [23]. The Spider has been validated in adolescents (aged 13 to 18) [23] and adults (aged 18 to 65) [24]. Moderate to strong convergent validity was found in all eight domains with adolescents and adults [23, 24]. Known-group validity was established in all eight domains with adults [23], and seven domains with adolescents [24]. To support use of the Spider questionnaire as an outcome measure for screening and longitudinal follow-up, reliability analysis is required. This study aimed to investigate the test-retest reliability

and internal consistency of the Spider domains and total score for use in tracking changes over time.

Methods

A repeated-measures survey design was used to assess absolute and relative reliability and internal consistency of the Spider questionnaire. The Short-Form 36 (SF-36) [25] or the PEDS-QL Generic Core Scales (PEDS-QL GCS) [26] were used for adults and adolescents, respectively, to monitor response stability during test-retest analysis. The SF-36 is an extensively validated questionnaire, is frequently used in symptomatic hypermobility (SH) research [27, 28], and has been shown to be sensitive to change [29]. It has 36-items across eight scales which measure QoL in both physical and emotional domains. The PEDS-QL GCS [26] also has sound psychometric properties and has been used in SH research [30, 31].

Recruitment

Participants were recruited with the help of three hypermobility patient support organisations: The Ehlers-Danlos Society, Ehlers-Danlos Support UK and the Hypermobility Syndromes Association. Study advertisements were published on websites and social media accounts and distributed via email newsletters. Participants self-selected based on the inclusion criteria outlined in Table 2. A link to a survey hosted by REDCap®, an online survey platform, was included in the advertisements. A sample size calculation was conducted for the reliability analysis. To demonstrate reliability of greater than 0.50, a total sample size of 424 was advised, assuming an expected intraclass correlation coefficient (ICC) of 0.60, measurements at two timepoints, and an attrition rate of 10%. This was based on a literature search in which the most conservative estimate for reliability across the different dimensions of the Spider questionnaire was used. Based on this review, the expected ICC was set at 0.6 using reported values for subjective symptoms in a related musculoskeletal condition [32, 33].

Table 2: Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
• Aged between 13 and 65 • Diagnosis of HSD/hEDS • Diagnosis of symptomatic hypermobility • Able to understand and communicate in English	• Younger than 13 or older than 65 • Unrelated neurological comorbidities (for example cerebral palsy) • Musculoskeletal comorbidities unrelated to hypermobility (for example a traumatic injury) • Unrelated rheumatological comorbidities (for example juvenile idiopathic arthritis) • GHJ with no symptoms

Key: *hEDS* – hypermobile Ehlers Danlos Syndrome, *HSD* – Hypermobility Spectrum Disorder, *GHJ* – Generalised Joint Hypermobility

Procedure

Potential participants accessed the online survey link, where they could read an information sheet and provide consent before proceeding to the questionnaires. Informed consent was obtained from all participants; for adolescents under 16 years old, consent was also gained from their parent/guardian. Consented participants were asked to provide basic demographic data including age, sex assigned at birth, gender, ethnicity, geographical location, diagnosis and the clinician who made the diagnosis. They were also asked to provide an email address for automated repeat survey reminders.

Participants were asked to complete two surveys one-to-two weeks apart. To ensure accurate comparison between surveys for test-retest reliability, sufficient time is required between tests to reduce recall bias, but not so much time that symptoms are likely to change [34]. An interval of one-to-two weeks is considered appropriate and is commonly used for test-retest studies; therefore, this interval was used in the present study [35]. Both surveys contained the Spider questionnaire and either the SF-36 or PEDS-QL GCS for adults and adolescents respectively.

Data analysis

IBM® SPSS® Statistics (version 29) was used for descriptive and inferential statistics analyses. Data were stored on the UCL Data Safe Haven, and only anonymised data were transferred out. Only completed questionnaires were analysed; partially completed questionnaires were excluded. Normality was assessed through visual inspection of histograms and the Shapiro-Wilk test. The Spider scores were calculated for each domain (mean item score, with a maximum score of 100) and as a total score (mean of the domain scores, with a maximum score of 100). Higher scores reflect a greater perceived impact of the symptoms on daily life. SF-36 scores were calculated using the scoring guidelines, and the Physical Component Score (PCS) and the Mental Component Score (MCS) were calculated using the recommended algorithm [36]. PEDS-QL GCS scores were calculated using the scoring guidelines. The results for the SF-36 and PEDS-QL are scored in the opposite direction to the Spider, with higher results indicating better QoL.

Analysis included both relative and absolute reliability to provide a comprehensive assessment of the Spider's reliability. Relative reliability, typically measured using ICC, provides information on a test's ability to consistently differentiate between participants and maintain their relative position in the group. Test-retest reliability of the Spider was calculated using data from timepoint one (TP1) and timepoint two (TP2) using ICC analysis. A two-way mixed-effects, single measure, absolute agreement model was used. ICC values between 0.50 and 0.75 were interpreted as moderate reliability, values between 0.75 and 0.90 as good reliability, and values greater than 0.90 as excellent reliability [35]. Pearson's correlation coefficients between the Spider total score and the SF-36/PEDS-QL were calculated at TP1 and TP2 to monitor stability.

Absolute reliability, which reflects the reliability of scores for individuals measured at different times, was assessed using the Standard Error of Measurement (SEM), Minimal Detectable Change (MDC) and 95% limits of agreement (LoA) from Bland-Altman analysis. The SEM was

used to assess within-participant variability and was calculated using a pooled standard deviation (SD) from both TP1 and TP2 with the formula $SD_{pooled} \times \sqrt{}(1 - ICC)$ [37]. The MDC was calculated to quantify the smallest change required to indicate a change beyond measurement error using the formula SEM x 1.96 x $\sqrt{}2$ [38]. Finally, the 95% LoA via Bland-Altman plots were used to assess agreement between TP1 and TP2, with narrower LoA indicating greater agreement [39]. These were calculated using the mean difference between timepoints ±1.96 SD. The 95% confidence interval (CI) around the mean difference was also calculated to examine systematic bias [40].

Internal consistency (the degree to which questionnaire items measure the same concept and are interrelated) was calculated using Cronbach's alpha (α) coefficient from data collected at TP1. An α of 0.70 to 0.95 is considered acceptable [41].

Results

Participants

In total, 1020 participants accessed the survey. At TP1, 800 individuals participated. Of this cohort, 617 participants completed surveys at both TP1 and TP2 and were included in analysis (Table 3). Participants were predominantly female (95.8%), White (94.8%) and from the United Kingdom (UK; 79.3%). The mean participant age was 38.3 ± 13.4 years. The majority were diagnosed with hEDS (71.3%), 26.1% had HSD and 2.6% had SH. Most participants were diagnosed by a rheumatologist (58.3%).

Table 3: Participant characteristics

Participant characteristics		Participants included in analysis (n=617)
Age (years)†		38.3±13.4
Sex*	Female	591 (95.8%)

	Intersex	1 (0.2%)
	Male	25 (4.1%)
Gender*	Woman	541 (87.7%)
	Man	34 (5.5%)
	Non-binary	33 (5.3%)
	A gender not listed here	1 (0.2%)
	Unsure	5 (0.8%)
	Prefer not to say	3 (0.5%)
Gender same as sex*	Yes	566 (91.7%)
	No	43 (7.0%)
	Prefer not to say	8 (1.3%)
Ethnicity*	Asian/ Asian British	2 (0.3%)
	Black/ African/ Caribbean/ Black British	2 (0.3%)
	Mixed/ Multiple Ethnicity	24 (3.9%)
	Others	4 (0.7%)
	White	585 (94.8%)
Country of residence*	Africa	1 (0.2%)
	Asia	1 (0.2%)
	Australia/ Oceania	37 (6.0%)
	Europe	26 (4.2%)
	North America and Canada	57 (9.2%)
	Republic of Ireland	6 (1.0%)
	South America	0
	United Kingdom	489 (79.3%)
Diagnosis*	HSD	161 (26.1%)
	hEDS	440 (71.3%)
	SH	16 (2.6%)
Diagnosed by*	Rheumatologist	360 (58.3%)
	Geneticist	98 (15.9%)
	Paediatrician	14 (2.3%)
	GP	62 (5.7%)
	Physiotherapist	35 (5.7%)
	Other	48 (7.8%)

Key: † - numeric data are presented as mean± standard deviation. * Categorical data are presented as counts and frequencies in percentages (%). *GP* – general practitioner, *hEDS* – hypermobile Ehlers Danlos Syndrome, *HSD* – Hypermobility Spectrum Disorder, *SH* – Symptomatic Hypermobility.

Reliability

Relative reliability of the Spider was measured through ICC (Table 4). ICC analysis showed good test-retest reliability between timepoints in all eight domains (0.79 to 0.88) and excellent reliability

for the total score (0.92). Correlations between the SF36/PEDS-QL and Spider total scores were very similar at TP1 and TP2. Correlations between the Spider total score and SF-36 PCS were r=-0.47 and -0.48 at TP1 and TP2, respectively, and correlations with the SF-36 MCS were r=-0.44 and -0.46, respectively. Correlations between the Spider total score and PEDS-QL total score were r=-0.68 and -0.70 at TP1 and TP2, respectively. Differences of only r=0.01 to 0.02 between timepoints suggest minimal change and therefore temporal stability. The mean difference between scores at TP1 and TP2 was close to zero for six domains and the total score. The gastrointestinal and anxiety domains had slightly larger mean score differences (2.10 and 1.97, respectively), although these remained small relative to the maximum domain score of 100. The 95% CIs around the mean differences for the pain, neuromusculoskeletal, urogenital, cardiac dysautonomia and depression domains included zero, suggesting no evidence of systematic bias between TP1 and TP2. The 95% CI around the mean differences for the fatigue, gastrointestinal and anxiety domains and the total score did not include zero and reflected small increases in scores between TP1 and TP2, suggesting a trend towards higher scores on retesting. SEM values for the domains and total score ranged from 4.0 to 11.60 and MDC values ranged from 11.08 to 32.13. The 95% LoA varied between domains, from -19% to +33% of the total domain score, and -11% to +13% of the maximum total questionnaire score, indicating some variation over time (Figure 1).

Internal consistency

Seven of the eight domains demonstrated acceptable internal consistency (α=0.70 to 0.87). An α of 0.69 was found for the pain domain, 0.01 below the recommended threshold of 0.70. The total score also demonstrated acceptable internal consistency (α=0.91) (Table 4).

Discussion

This study aimed to add to the existing body of work analysing the psychometric properties of the Spider. Construct validity has been established [23, 24], and the present study investigated the

reliability and internal consistency of the Spider individual domains and total score. These psychometric properties are discussed in further detail below. In summary, the results demonstrate good to excellent test-retest reliability based on ICC. SEM values for the symptom domains were relatively low, suggesting reasonable stability between timepoints. Internal consistency was acceptable for seven of the eight domains and the total score.

Reliability

Relative reliability (test-retest reliability) was assessed to determine the Spider's ability to produce similar scores on two separate occasions when the condition being measured is assumed to have remained stable [42]. ICC analysis demonstrated good to excellent reliability across all domains and the total score between time points (0.79 to 0.92). Five domains showed good test-retest reliability (pain, fatigue, neuromusculoskeletal, gastrointestinal and anxiety), and three domains (urogenital, cardiac dysautonomia and depression) and the total score showed excellent test-retest reliability. The lower bounds of the 95% CIs for each domain and the total score also remained at or above 0.75, which is considered good. Absolute reliability was assessed through SEM and MDC. SEM values were relatively low, generally below 10% of the maximum possible score of 100, with the exception of the anxiety and depression domains (11.60 and 10.69, respectively). This suggests reasonable stability between timepoints, with somewhat greater variation in mental health domains, which may be expected given their reported temporal variability [35]. Stability between timepoints is also reflected in the consistency of the correlations between the Spider scores and the SF-36/PEDS-QL, which remained very similar at both TP1 and TP2. MDC values were substantially larger than SEM values (11.08 to 32.13); this may partly reflect the increments used in the Spider's Likert scoring. During score calculation, the Likert scale is converted to 25-point increments from 0 to 100 (no symptoms/no impact, mild, moderate, severe, disabling), so a change beyond measurement error may be reflected in a relatively large score change. Bland-Altman analysis provided a method for assessing agreement between

scores at each timepoint and variation over time. This demonstrated minimal bias in most domains and total score, with small positive mean differences on retesting in the gastrointestinal and anxiety domains, indicating slightly higher scores (and therefore greater symptom impact) at TP2. The LoA were relatively wide, reflecting a variation of approximately one Likert-scale category in both directions across domains. The total score showed less variation (-10.61 to 12.51) which may be due to the larger number of questions contributing to the total score compared with the individual domains, which contain between three and five questions.

Internal consistency of the individual domains and total score was assessed to identify the extent to which items within each scale measure the intended concept cohesively [43]. Seven of the eight domains and the total score showed acceptable internal consistency (α=0.70 to 0.91). A Cronbach's alpha of α = 0.69 was observed for the pain domain. This result is close to the conventional threshold of 0.70 and may reflect the breadth of the pain construct assessed by this domain. The pain domain assesses the presence and impact of joint pain, widespread pain, headaches or migraines, and allodynia/hyperalgesia across four questions. As discussed, the prevalence of pain is high in this population, although the prevalence of different types of pain varies. In people with HSD/hEDS, 90% report generalised pain [6], 83.2% report specific joint pain [8], and 60.5% and 42.5% report headaches and migraines, respectively [8]. This heterogeneity of pain presentations, assessed using only four questions, may have contributed to the slightly lower internal consistency of this domain. The remaining domains and the total score demonstrated acceptable internal consistency without very high coefficients (>0.95) which could indicate item redundancy, allowing the Spider to retain its brevity.

<u>Strengths and Limitations:</u>

This study used several robust methods to analyse reliability, ensuring that both absolute and relative reliability were assessed. The use of an additional measure (the SF-36 or PEDS-QL)

provided further confidence that there was relative stability between timepoints. The study sample was large and had international reach, with sufficient numbers for reliability assessment. As in previous Spider validity studies, the sample was predominantly White and female. Although SH is more prevalent in females, the proportion of females in this sample was disproportionately high compared with reported diagnostic ratios in people with HSD/hEDS [44]. The sample also lacked ethnic diversity, as observed in previous Spider validation studies. This may reflect the complex sociocultural factors which influence health-seeking behaviours, including access to support organisations and willingness to discuss health [45, 46]. The limited diversity reduces the generalisability of the findings to populations other than White females. Participants self-reported their diagnosis, and, although information about the clinician who made the diagnosis was collected, diagnoses were not independently corroborated. Future research should consider objective assessment using tools such as the Beighton Scale [47] and the five-point JH screening questionnaire [48]. The attrition rate was 16.6% between timepoints, within the reported acceptable rate of <20% [49]. Chi-squared or Fisher's exact tests were used to compare demographic variables between participants at TP1 and those retained at TP2. No significant differences were identified, suggesting no evidence of differential attrition on the measured demographic variables. The retained sample remained sufficiently large for the planned reliability analysis.

Developments and future research

A website has been created to enable clinicians, researchers and patients to download the Spider freely, at www.thespidertool.com. A future natural history study, using the Spider as a symptom profiling tool, will assess relationships between demographic variables and symptoms and explore possible trends and patterns. Ongoing research is assessing minimal clinically important differences (MCIDs) to support interpretation of the Spider as an outcome measure for tracking change over time. This will help clinicians understand when changes in Spider scores are

meaningful to patients. Translations of the Spider will increase its international accessibility, and work has begun translating it into Arabic, Dutch and Spanish. Further translations are recommended to improve accessibility. Future studies should aim to recruit more diverse populations using targeted recruitment methods. Objective diagnostic criteria, such as the Beighton scale and the five-part questionnaire should be used to support diagnostic classification, or participants could be recruited through registries in which diagnoses have been confirmed.www.thespidertool.com

Conclusion

This work adds to the series of studies establishing the psychometric properties of the Spider, a symptom impact tool for people with hypermobility related disorders. Relative and absolute reliability were assessed, demonstrating good to excellent test-retest reliability. Internal consistency was acceptable for seven of the eight domains and the total score; the pain domain was marginally below the conventional threshold.

List of abbreviations

CI: Confidence intervals

GJH: Generalised Joint Hypermobility

hEDS: Hypermobile Ehlers Danlos Syndrome

HSD: Hypermobility Spectrum Disorder

ICC: Intraclass correlation coefficient

LoA: Limits of Agreement

MCS: Mental Component Score

MDC: Minimal Detectable Change

NMSK: Neuromusculoskeletal

OR: Odds ratio

PCS: Physical Component Score

PEDS-QL GCS: PEDS-QL Generic Core Scales

POTS: Postural Orthostatic Tachycardia Syndrome

QoL: Quality of life

SEM: Standard Error of Measurement

SD: Standard Deviation

SF-36: Short Form 36

SH: Symptomatic Hypermobility

TP1: Timepoint One

TP2: Timepoint Two

UCL: University College London

References

1. Blajwajs L, Williams J, Timmons W, Sproule J. Hypermobility prevalence, measurements, and outcomes in childhood, adolescence, and emerging adulthood: a systematic review. *Rheumatol Int*. 2023;43(8):1423-44.

2. Nicholson LL, Simmonds J, Pacey V, De Wandele I, Rombaut L, Williams CM, et al. International Perspectives on Joint Hypermobility: A Synthesis of Current Science to Guide Clinical and Research Directions. *J Clin Rheumatol*. 2022;28(6):314-20.

3. Castori M. Deconstructing and reconstructing joint hypermobility on an evo-devo perspective. *Rheumatology (Oxford)*. 2021;60(6):2537-44.

4. Castori M, Tinkle B, Levy H, Grahame R, Malfait F, Hakim A. A framework for the classification of joint hypermobility and related conditions. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*. 2017;175(1):148-57.

5. Tofts LJ, Simmonds J, Schwartz SB, Richheimer RM, O'Connor C, Elias E, et al. Pediatric joint hypermobility: a diagnostic framework and narrative review. *Orphanet J Rare Dis*. 2023;18(1):104.

6. Chopra P, Tinkle B, Hamonet C, Brock I, Gompel A, Bulbena A, et al. Pain management in the Ehlers-Danlos syndromes. *Am J Med Genet C Semin Med Genet*. 2017;175(1):212-9.

7. 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.

8. Fairweather D, Bruno KA, Darakjian AA, Bruce BK, Gehin JM, Kotha A, et al. High overlap in patients diagnosed with hypermobile Ehlers-Danlos syndrome or hypermobile spectrum disorders with fibromyalgia and 40 self-reported symptoms and comorbidities. *Frontiers in Medicine*. 2023;10.

9. Ruiz Maya T, Fettig V, Mehta L, Gelb BD, Kontorovich AR. Dysautonomia in hypermobile Ehlers–Danlos syndrome and hypermobility spectrum disorders is associated with exercise

intolerance and cardiac atrophy. *American Journal of Medical Genetics Part A*. 2021;185(12):3754-61.

10. Thwaites PA, Gibson PR, Burgell RE. Hypermobile Ehlers-Danlos syndrome and disorders of the gastrointestinal tract: What the gastroenterologist needs to know. *J Gastroenterol Hepatol*. 2022;37(9):1693-709.

11. Tai FWD, Palsson OS, Lam CY, Whitehead WE, Sperber AD, Tornblom H, et al. Functional gastrointestinal disorders are increased in joint hypermobility-related disorders with concomitant postural orthostatic tachycardia syndrome. *Neurogastroenterology & Motility*. 2020;32(12):e13975.

12. Gazit Y, Nahir AM, Grahame R, Jacob G. Dysautonomia in the joint hypermobility syndrome. *Am J Med*. 2003;115(1):33-40.

13. Gilliam E, Hoffman JD, Yeh G. Urogenital and pelvic complications in the Ehlers-Danlos syndromes and associated hypermobility spectrum disorders: A scoping review. *Clin Genet*. 2020;97(1):168-78.

14. Smith TO, Easton V, Bacon H, Jerman E, Armon K, Poland F, et al. The relationship between benign joint hypermobility syndrome and psychological distress: a systematic review and meta-analysis. *Rheumatology*. 2014;53(1):114-22.

15. Pasquini M, Celletti C, Berardelli I, Roselli V, Mastroeni S, Castori M, et al. Unexpected association between joint hypermobility syndrome/Ehlers-Danlos syndrome hypermobility type and obsessive-compulsive personality disorder. *Rheumatol Int*. 2014;34(5):631-6.

16. Orenius T, Kautiainen H, Louhi M, Montin L, Bulbena A, Lindgren K-A. Health-Related Quality of Life and Psychological Distress in Patients With Hypermobility Type Ehlers-Danlos Syndrome. *SAGE Open*. 2022;12(2):21582440221091237.

17. Pacey V, Tofts L, Adams RD, Munns CF, Nicholson LL. Quality of life prediction in children with joint hypermobility syndrome. *J Paediatr Child Health*. 2015;51(7):689-95.

18. Berglund B, Pettersson C, Pigg M, Kristiansson P. Self-reported quality of life, anxiety and depression in individuals with Ehlers-Danlos syndrome (EDS): a questionnaire study. *BMC Musculoskeletal Disorders*. 2015;16(1):89.

19. Sahin E, Gulbahar S, Baydar M, Ozcan Soylev G, Bircan C, El O, et al. Impaired quality of life and functional status in patients with benign joint hypermobility syndrome. *MYOPAIN*. 2015;23(1-2):28-33.

20. De Wandele I, Rombaut L, Malfait F, De Backer T, De Paepe A, Calders P. Clinical heterogeneity in patients with the hypermobility type of Ehlers-Danlos syndrome. *Res Dev Disabil*. 2013;34(3):873-81.

21. Scheper MC, Pacey V, Rombaut L, Adams RD, Tofts L, Calders P, et al. Generalized Hyperalgesia in Children and Adults Diagnosed With Hypermobility Syndrome and Ehlers-Danlos Syndrome Hypermobility Type: A Discriminative Analysis. *Arthritis Care Res (Hoboken)*. 2017;69(3):421-9.

22. National Academies of Sciences EaM. In: Wedge RA, Cartaxo T, Spicer CM, Volberding PA, eds. *Selected Heritable Disorders of Connective Tissue and Disability*. Washington (DC): National Academies Press
; 2022.

23. Ewer E, Kazkaz H, Ninis N, Rowe P, De Pauw R, Tang E, et al. The Spider; a multisystemic symptom impact tool for people with hypermobility related disorders. Initial validation in adolescents. *The Journal of Pediatrics: Clinical Practice*. 2024:200098.

24. Ewer ER, De Pauw, R., Kazkazk, H. et al. The Spider: a visual, multisystemic symptom impact questionnaire for people with hypermobility-related disorders—validation in adults. *Clin Rheumatol* 2024.

25. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. *Health Econ*. 1993;2(3):217-27.

26. Varni JW, Seid M, Knight TS, Uzark K, Szer IS. The PedsQL 4.0 Generic Core Scales: sensitivity, responsiveness, and impact on clinical decision-making. *J Behav Med*. 2002;25(2):175-93.

27. Aukštuolytė-Bačienė E, Daunoravičienė A, Tamulionytė V, Berškienė K, Narbutaitė J, Razon S, et al. Present but Ignored: Physical Condition and Health-Related Quality of Life in College-Aged Females with Generalized Joint Hypermobility. *Healthcare*. 2024;12(11):1065.

28. Molander P, Novo M, Ringqvist Å, Hållstam A, Hesser H, Löfgren M, et al. Interdisciplinary pain rehabilitation for patients with Ehlers-Danlos syndrome and hypermobility spectrum disorders. *J Rehabil Med*. 2024;56:jrm12431.

29. Brazier JE, Harper R, Jones NMB, O'Cathain A, Thomas KJ, Usherwood T, et al. Validating The SF-36 Health Survey Questionnaire: New Outcome Measure For Primary Care. *BMJ: British Medical Journal*. 1992;305(6846):160-4.

30. Mu W, Muriello M, Clemens JL, Wang Y, Smith CH, Tran PT, 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(4):561-9.

31. Maarj M, Pacey V, Tofts L, Clapham M, Coda A. The Impact of Podiatric Intervention on the Quality of Life and Pain in Children and Adolescents with Hypermobility. *International Journal of Environmental Research and Public Health*. 2023;20(17):6623.

32. Ford JJ, Story I, McMeeken J. The test-retest reliability and concurrent validity of the Subjective Complaints Questionnaire for low back pain. *Man Ther*. 2009;14(3):283-91.

33. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15(2):155-63.

34. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin Trials*. 1991;12(4 Suppl):142s-58s.

35. Polit DF. Getting serious about test–retest reliability: a critique of retest research and some recommendations. *Quality of Life Research*. 2014;23(6):1713-20.
36. Ware JE. SF-36 physical and mental health summary scales : a user's manual. 1994.
37. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Phys Ther*. 2006;86(5):735-43.
38. Schuck P, Zwingmann C. The 'smallest real difference' as a measure of sensitivity to change: a critical analysis. *Int J Rehabil Res*. 2003;26(2):85-91.
39. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-10.
40. Chuang L-L, Lin K-c, Hsu A-L, Wu C-Y, Chang K-C, Li Y-C, et al. Reliability and validity of a vertical numerical rating scale supplemented with a faces rating scale in measuring fatigue after stroke. *Health and quality of life outcomes*. 2015;13:91.
41. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011;2:53-5.
42. Portney L, Watkins MP. Foundation of Clinical Research. Application to Practice. *Upper Saddle River*. 2009:61-77.
43. Shou Y, Sellbom M, Chen H-F. 4.02 - Fundamentals of Measurement in Clinical Psychology. In: Asmundson GJG, ed. *Comprehensive Clinical Psychology (Second Edition)*. Oxford: Elsevier; 2022. p. 13-35.
44. Palmer S, Manns S, Cramp F, Lewis R, Clark EM. Test-retest reliability and smallest detectable change of the Bristol Impact of Hypermobility (BIoH) questionnaire. *Musculoskeletal Science and Practice*. 2017;32:64-9.
45. Memon A, Taylor K, Mohebati LM, Sundin J, Cooper M, Scanlon T, et al. Perceived barriers to accessing mental health services among black and minority ethnic (BME) communities: a qualitative study in Southeast England. *BMJ Open*. 2016;6(11):e012337.
46. Ayalon L, Young MA. Racial Group Differences in Help-Seeking Behaviors. *The Journal of Social Psychology*. 2005;145(4):391-403.

47. Beighton P, Solomon L, Soskolne CL. Articular mobility in an African population. *Ann Rheum Dis*. 1973;32(5):413-8.

48. Hakim AJ, Grahame R. A simple questionnaire to detect hypermobility: an adjunct to the assessment of patients with diffuse musculoskeletal pain. *Int J Clin Pract*. 2003;57(3):163-6.

49. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *The Lancet*. 2002;359(9308):781-5

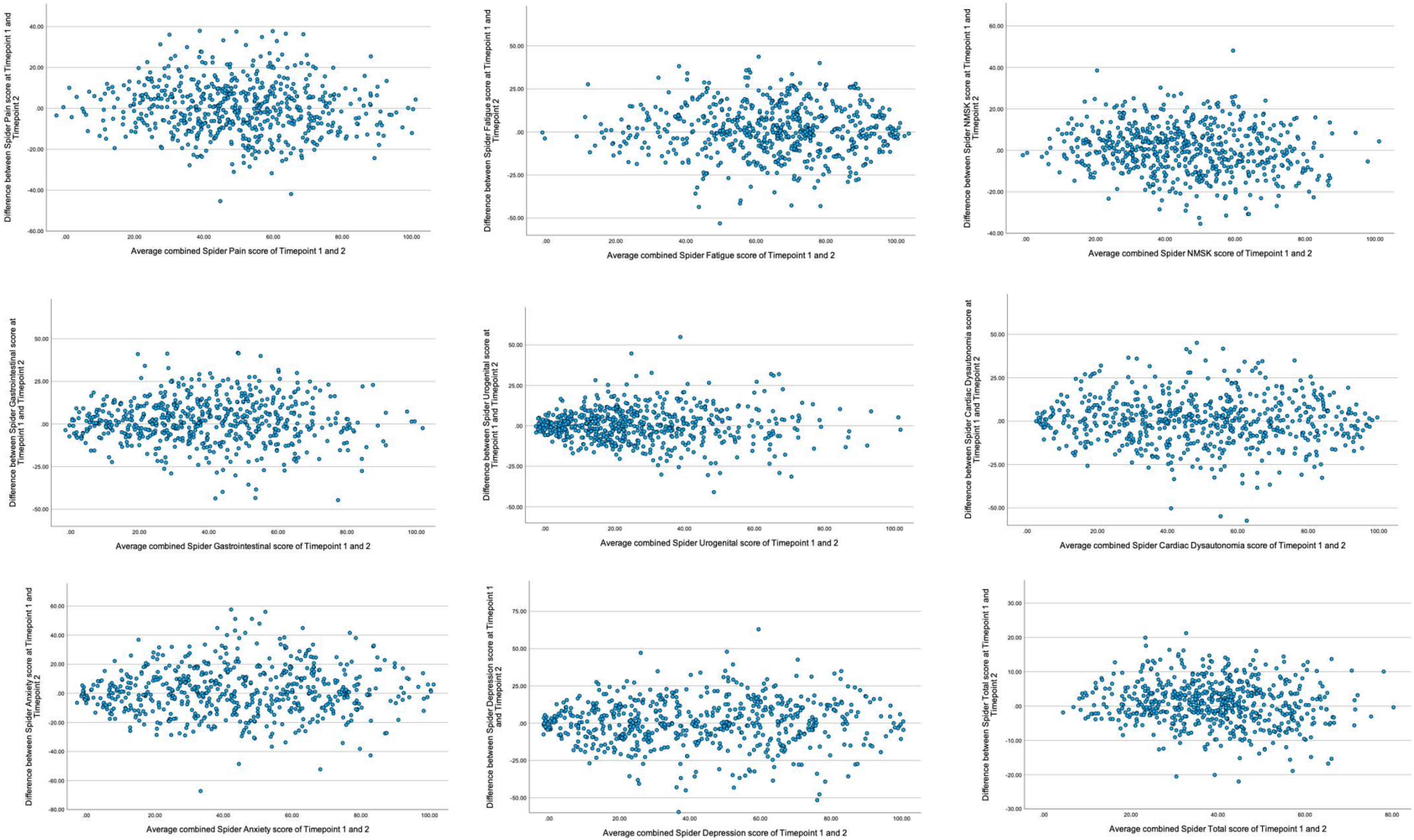


Fig. 1 Bland Altman plots of Spider domains and total score. Plots have been jittered to improve visualisation of data distribution.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table4.docx