RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics PART A

Lack of Consensus on Tests and Criteria for Generalized Joint Hypermobility, Ehlers–Danlos Syndrome: Hypermobile Type and Joint Hypermobility Syndrome

Lars Remvig,[1]* Lise Flycht,[2] Karl B. Christensen,[3] and Birgit Juul-Kristensen[4]

[1]Department of Infectious Medicine and Rheumatology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

[2]IT-Department, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

[3]Department of Biostatistics, University of Copenhagen, Copenhagen, Denmark

[4]Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

Manuscript Received: 8 March 2013; Manuscript Accepted: 27 November 2013

The objectives of this study were to register clinicians performance and opinion of importance of clinical tests for generalized joint hypermobility (GJH), Ehlers–Danlos syndrome, hypermobility type (EDS-HT) and joint hypermobility syndrome (JHS), and to reach a consensus among clinicians on criteria for diagnosing GJH, EDS-HT and JHS. A panel of clinicians answered questions about how to perform and interpret clinical tests and rated test importance on an 11-box scale. The questionnaire was developed on the basis of information from focus groups and the literature. Cronbach's α was used as a measure of internal consistency/consensus among the panelists. The results showed Cronbach's α on importance score of items for diagnosing GJH, EDS-HT and JHS was 0.61, 0.79, and 0.44, respectively. Panelist-group correlation for the three conditions varied substantially (−0.46 to 0.89, 0.03 to 0.68, and −0.07 to 0.68) indicating heterogeneity among the panelists. There was agreement on which tests to use, but performance of the tests (i.e., the specific maneuvers) varied considerably inclusive use of tests with unknown reliability. Furthermore, agreement on the diagnostic criteria varied. We conclude that the level of consensus for the importance of various items for diagnosing GJH, EDS-HT and JHS, was below the required limit (Cronbach's α >0.90) for clinical decision-making and diagnosing. Consensus on tests and criteria through a Delphi process could not be reached. Better descriptions of, and reliability studies on, test maneuvers and criteria sets for these conditions are needed. Subsequent intensive training and implementation of these tests and criteria, nationally as well as internationally should be established.

Key words: consensus; Cronbach's α; text variability; Ehlers–Danlos syndrome; joint hypermobility syndrome

How to Cite this Article:
Remvig L, Flycht L, Christensen KB, Juul-Kristensen B. 2014. Lack of consensus on tests and criteria for generalized joint hypermobility, Ehlers–Danlos syndrome: Hypermobile type and joint hypermobility syndrome.
Am J Med Genet Part A 164A:591–596.

INTRODUCTION

Generalized joint hypermobility (GJH) is part of the diagnostic criteria for a wide range of heritable connective tissue disorders (HCTD). Among them joint hypermobility syndrome (JHS) [Grahame et al., 2000] and Ehlers–Danlos syndrome, hypermobility type (EDS-HT) [Beighton et al., 1998]—possibly one and the same condition [Remvig et al., 2011]—are probably the most common; but as is the case in most HCTD, JHS/EDS-HT does not have one or more additional specific clinical feature and in addition cannot be diagnosed through laboratory tests. Therefore, standardizing clinical diagnostic criteria must be stringent in order to differentiating JHS/EDS-HT both from asymptomatic GJH and from rarer HCTD. In other words, better definitions of the conditions are needed for two reasons: (i) for better stratifying the general population between individuals at risk of developing GJH related symptoms and individuals not at risk, and (ii) for identifying an efficient tool guiding the choice between performing or not molecular testing and adequately selecting patients for research.

Conflict of interest: None.
*Correspondence to:
Lars Remvig, MD, DMSc, Retired, Gassehaven 56, 2840 Holte, Denmark.
E-mail: l.remvig@dadlnet.dk
Article first published online in Wiley Online Library (wileyonlinelibrary.com): 24 January 2014
DOI 10.1002/ajmg.a.36402

Furthermore, it is striking that neither the two above mentioned sets of criteria nor the literature describe precisely how to perform the clinical tests used—being joint tests, skin tests or specific symptoms in the patient history—to make the clinical diagnosis (or diagnoses). Consequently, considerable variation in test performance is presented in scientific publications, in textbooks and on the web, which together with a variation in the definition of GJH [Remvig et al., 2007] further increases the variation in diagnostic results for these disease entities.

These problems call for an agreement among international opinion-makers on how to perform the clinical tests, as well as a consensus on the criteria for GJH, for JHS and for EDS-HT. The Delphi method [Dalkey and Helmer, 1963; Dalkey, 1969; Dalkey et al., 1969], which is an attractive method to improve the quality of judgements of relatively uncertain issues [Milholland et al., 1973], has been accepted as a method to achieve consensus on controversial subjects within health science [Fink et al., 1984] and recently been used to establish consensus for diagnostic criteria [Graham et al., 2003; John, 2010].

The purposes of our study, using the Delphi method, were (i) to describe how international specialists in the field perform the clinical tests for the conditions GJH, EDS-HT and JHS, and whether they relate gender, age, and race to the criteria for these conditions, (ii) to measure the level of agreement amongst these specialists for the clinical tests and criteria using Cronbach's α as a measure of the internal consistency of the group, and (iii) to reach an international consensus on a proposal for criteria of EDS-HT and JHS.

MATERIALS AND METHODS

The Delphi technique includes the following phases.

Forming an Investigator Group

The authors, Lars Remvig, Birgit Juul-Kristensen, and Lise Flycht, formed the investigator group (IG) who wrote the protocol for the Delphi Study, reviewed the literature with respect to criteria items, formed the focus groups, collected and condensed suggestions for criteria items given by the focus groups, and were obliged to accept all items recommended. The IG also formed the panel by inviting panelists confidentially. Members of the IG could not be members of the focus groups or the panel.

Forming Focus Groups

Participants at the third Nordic Meeting in Hypermobility Research, Copenhagen 2009 formed three focus groups with four participants in each group, mixing nationalities, academic education, and medical specialization. The focus groups had (i) to suggest criteria items for diagnosing GJH, JHS and EDS-HT (clinical tests, cut-off levels, etc.) (ii) to discuss variations depending on gender, age, and race, and (iii) to discuss and propose principles for selection/recruitment of members for the panel.

Forming a Panel

From a PubMed search with the searching term "hypermobility," the IG screened 525 papers published between May 1, 2005 and April 30, 2010 for establishing the panel. Only 150 of these dealt with joint hypermobility, and were screened thoroughly by the IG in order to exclude reviews, case studies, letters, viewpoints, editorials, retrospective studies, question-based studies, and guidelines. Further, publications without information on e-mail address for correspondence were excluded. Only the latest publication from an author or a department was included, ending up with 44 relevant publications.

From each publication the IG selected either: the first author; or the one to whom correspondence should be addressed, or the one who was responsible for the examination and diagnosis of the patients. In three cases, the authors could not be addressed by e-mail, leaving 41 to be invited as panelists. Fifteen of the 41 authors responded positively being dermatologists, geneticists, pediatricians, physical medicine specialists, rheumatologists, physiotherapists, and one psychologist, coming from 11 different countries in the western world. They all accepted to take part in a second questionnaire round. The composition of the panel was not disclosed to the panelists.

Forming a Questionnaire

Based on the proposals of the focus groups and the literature, the IG compiled a list of items (Appendices 1 and 2—see supporting information online) for the questionnaire. Nine main items were related to GJH, 16 were related to EDS-HT, and one to the opinion of the panelists, whether EDS-HT and JHS are two different disease entities. If respondents answered in the affirmative, the very same questions as those posed for the EDS-HT were asked again, but this time for JHS. For each item, panelists had to answer the following question: "How important/relevant to you is [the specific item] in diagnosing GJH, EDS-HT/JHS respectively?" An 11-box visual analogue scale with the anchors: "Completely unimportant/not relevant (0)" and "Extremely important/relevant (10)" was used to record the responses, and the median ± range for the group for each item comprised an "Importance score." Some questions gave rise only to dichotomous answers, quantified as 0 or 10. Most of the items had sub-questions: for example, "skin signs" was divided into skin extensibility, scarring, bruising, etc.; skin extensibility was further divided into the location used to judge the sign (back of hand, back of forearm, volar aspect of forearm, etc.), how to perform the test (pinching the cutis, pinching the cutis + subcutis, in the sagittal or the horizontal plane, etc.), and when extensibility is pathologic ($\geq$1, $\geq$2, $\geq$3, $\geq$4, or $\geq$5 cm).

The questionnaire was converted into a web-based questionnaire using Enalyzer (www.enalyzer.com), tested and revised by the IG before being forwarded to the panelists by the IT-department, Rigshospitalet, including a letter with username and password. The answered questionnaire was returned to the IT-department, securing the anonymity of the panelists. When all responses were collected, the questionnaire was re-forwarded to the panelists, inclusive the computed results of each item (mean ± SD), and a request to re-consider/-evaluate their previous answers.

Statistics

Cronbach's α was used as the statistical index to estimate reliability of the sum of panelists' importance scores [Bland and Altman, 1997;

Graham et al., 2003]. In cases where the responses are highly correlated, they are considered to have high internal consistency, indicating a high level of consensus. The Cronbach's α value is recommended to be above 0.90 as a minimum for a diagnostic scale to be useful in clinical practice [Bland and Altman, 1997], and consequently an α-value >0.90 was a priori accepted as indicative of acceptable consensus. Also, the correlation between the various importance scores for each panelist and the total sum of the various importance scores of the other panelists was computed.

RESULTS

All 15 panelists responded to the first questionnaire after two reminders, and seven responded to the second after three reminders. Only data from round one is presented, due to the low response rate in the second round.

Generalized Joint Hypermobility

Cronbach's α for panelists' importance score of items to diagnose GJH was 0.61, and the individual panelist to group correlation ranged from −0.46 to 0.89. The median importance score for the Beighton tests varied from 7.0 (3–10) for the 5th finger test, to 9.0 (6–10) for the knee extension test. There was an almost 50/50 distribution among the panelists using only current or historical findings. Most of the panelists (13/15, i.e., 86.7%) used the Beighton 9-point scoring system; however cut-off levels varied for diagnosing GJH in adults, being either ≥4/9 (46%), ≥5/9 (46%) or ≥6/9 (8%). Different maneuvers were preferred for the performance of 4 of the 5 Beighton tests (Table I). The importance score for having the factors age, gender, and ethnicity influencing the cut-off level for GJH was 8.0 (range 0–10). Most of the panelists (12/13, 92%) were in favor of a change in the Beighton score cut-off level, rather than a change in ROM for this influence, and there was a clear tendency towards a negative correlation between age and cut-off level (Table II). Nine (75%) panelists suggested to increase the cut-off level for females, and five (42%) to decrease it for males. An increased cut-off level for certain ethnic groups was supported by 75%, with 25% advocating a decreased cut-off level for Caucasians. The importance score for including additional joint tests than the Beighton tests for diagnosing GJH was 8.0 (0–10), and when demanding presence of positive tests in upper as well as lower extremity joints the score was 6.0 (0–10).

Ehlers–Danlos Syndrome, Hypermobility Type

The answers on importance score in items for diagnosing EDS-HT resulted in Cronbach's α of 0.79. The individual panelists to group correlation ranged between 0.03 and 0.68 (Table III). The five most important criteria were: GJH, joint dislocations/subluxations, skin signs—particularly increased skin extensibility, soft tissue rheumatism, and arthralgia (Table IV). Thirteen (87%) of the panelists indicated dislocation(s)/subluxation(s) to have some degree of importance for diagnosing EDS-HT, but frequency and number of joints necessary to declare the sign abnormal varied considerably (Table V). When testing skin extensibility nine (60%) panelists advocated the back of the hand, and seven (47%) the volar aspect of the forearm (Table VI). Eleven (73%) panelists recommended both right and left side to be tested, and most panelists judged the cut-off level for abnormal skin extensibility to be >2 cm (ranging from >1 to >4 cm). Thirteen (87%) panelists found that atrophic, broad scarring had importance for diagnosing EDS-HT, with considerable variation in level of importance. Among these panelists, seven (54%) accepted a scar >5 mm as pathologic, six (46%) accepted *one* location as sufficient to judge scarring as being abnormal. In order to accept skin signs as a positive criterion sign for EDS-HT, seven

TABLE I. Number of Panelists, n (%), With Respect to Preferred Test Performance of 4 of the 5 Beighton Tests for Diagnosing Generalized Joint Hypermobility

Beighton test	Panelists' preferred test performance		
Thumb to volar aspect of forearm	With flexed elbow	With extended elbow and pronated hand	With extended elbow and supinated hand
5th finger dorsiflexion 90°	8 (53) With forearm and wrist flat on the table	2 (13) With forearm unsupported and wrist supported	5 (33) With forearm and wrist unsupported
Elbow extension >10°	8 (53) With arm flexed at the shoulder and hand supinated	1 (7) With arm abducted at the shoulder and hand supinated	6 (40) With arm hanging down and hand supinated
Knee extension >10°	8 (53) In standing position, foot on the floor	3 (20) In supine position on the couch	4 (27)
	8 (53)	7 (47)	

Note: Panelists were not asked about performance of Forward Bending.

TABLE II. Distribution of Panelists (%) With Respect to Their Perception of the Necessary and Sufficient Number of Positive Beighton Tests (0–9) for Diagnosing Generalized Joint Hypermobility in Relation to Age (n = 12)

Age in years	≥1/9	≥2/9	≥3/9	≥4/9	≥5/9	≥6/9
<13	—	—	—	1 (8)	2 (17)	9 (75)
13–19	—	—	—	2 (17)	7 (58)	3 (25)
20–50	—	—	1 (8)	6 (50)	5 (42)	—
>50	—	2 (17)	5 (41)	3 (25)	2 (17)	—

(47%) panelists demanded at least two abnormal skin signs; ten (67%) required abnormal skin extensibility to be present, and seven (47%) that atrophic, broad scarring was present.

Major and Minor Criterion Signs

More than 50% of panelists recommended GJH (100%), skin signs (71%), dislocation(s)/subluxation(s) (64%), and arthralgia (57%) as major signs. Recommended as minor criterion signs were eye signs (79%), gastrointestinal signs (79%), soft tissue rheumatism (79%), dysautonomy signs (64%), varicose veins, etc. (57%) and dental signs (57%). The presence of a minimum of two major signs was mandatory according to eight (57%) panelists, three major signs according to three (21%) panelists and four major signs according to one panelist. Among those requiring a minimum of two major signs, three panelists also recommended that these two major signs were sufficient to fulfill the criteria set for EDS-HT. The remaining five panelists suggested the presence of two to five minor signs.

Differentiation Between EDS-HT and JHS

Five panelists (33%) were of the opinion that EDS-HT and JHS phenotypic represent two different disease entities.

TABLE III. Individual Panelist-Group Correlation for Diagnosing Ehlers–Danlos Syndrome, Hypermobility Type, in the First Questionnaire Response in Relation to the Panelists' Education/Specialization

Panelist	r
Dermatologist	0.37
Geneticist	0.19
Pediatrician	0.58
Physical medicine 1	0.39
Physical medicine 2	0.03
Psychologist	0.23
Psychiatrist	0.24
Physiotherapist 1	0.61
Physiotherapist 2	0.57
Physiotherapist 3	0.68
Rheumatologist 1	0.26
Rheumatologist 2	0.49
Rheumatologist 3	0.60
Rheumatologist 4	0.26
Rheumatologist 5	0.57

TABLE IV. Panelists' Perception (Scale 0–10) on Various Clinical Signs as Criteria for Ehlers–Danlos Syndrome Hypermobility Type (EDS-HT) (n = 15) and Joint Hypermobility Syndrome (JHS) (n = 5)

Criterion	EDS-HT	JHS
Generalized joint hypermobility as defined	9.0 (5–10)	9.0 (5–10)
Joint dislocations and/or subluxations	8.0(0–10)	3.0 (0–10)
Skin signs	7.0 (1–10)	0.0 (0–8)
Increased extensibility	8.0 (1–10)	[a]
Atrophic broad scarring	7.0 (0–10)	[a]
Striae distensae	6.0 (0–10)	[a]
Soft velvety skin	6.0 (1–10)	[a]
Reduced skin thickness	5.0 (0–8)	[a]
Arthralgia	7.0 (0–10)	8.0 (0–10)
Soft tissue rheumatism	7.0 (0–10)	5.0 (0–10)
Dysautonomy signs	7.0 (0–10)	0.0 (0–10)
Physical dysfunction	7.0 (0–10)	0.0 (0–10)
Gastrointestinal tract sign	6.5 (0–9)	0.0 (0–8)
Varicose veins, hernias, rectal/uterine prolapse	6.0 (0–10)	0.0 (0–10)
Vascular signs	6.0 (0–10)	0.0 (0–8)
Local anesthesia effect	6.0 (0–9)	0.0 (0–6)
Psychiatric disorder(s)	5.0 (0–10)	0.0 (0–10)
Dental signs	5.0 (0–10)	0.0 (0–6)
Eye signs	5.0 (0–10)	0.0 (0–8)
Marfanoid habitus	5.0 (1–10)	0.0 (0–10)
Gorlin sign	4.0 (1–10)	0.0 (0–10)

Results are given as "importance score," median (range).
[a]Insufficient data (n = 2).

Joint Hypermobility Syndrome

The answers on importance score for items diagnosing JHS resulted in Cronbach's α of 0.44, and the panelist to group correlation varied between −0.72 and 0.68. GJH and arthralgia were judged as the most important signs in diagnosing JHS (Table IV), with soft tissue rheumatism and dislocation(s)/subluxation(s) having less importance. Three panelists recommended GJH to be a major criterion,

TABLE V. Panelists' Perception of the Necessary Combination of Frequency and Number of Joints With Dislocation (s) and/or Subluxation (s) to to Accept the Dislocation/Subluxation sign as a Positive Criterion Sign in Diagnosing Ehlers–Danlos Syndrome, Hypermobility Type

	Time				
Number	1 time	2 times	3 or more times	Missing values	Total
1 joint	0 (0)	8 (62)	4 (31)	1 (8)	13 (100)
2 joints	6 (46)	5 (38)	1 (8)	1 (8)	13 (100)
3 or more joints	8 (62)	1 (8)	2 (15)	2 (15)	13 (100)

Results are given as number of panelists, n (%).

TABLE VI. Panelists' Perception of How and Where to Judge Skin Extensibility and of the Cut-Off Level for Abnormality (n = 15)

Panelist no	Pinching	Volar forearm	Dorsal hand	Chest
1	Cutis + subc	>2 cm[a]		+
2	Cutis + subc	>2 cm[a]	+	
3	Cutis		>3 cm	
4	Cutis		>2 cm	
5	Cutis	>3 cm		
6	Cutis		>2 cm	
7	Cutis	>4 cm		
8	Cutis + subc		>2 cm	
9	Cutis	>2 cm		
10	Cutis		>2 cm	
11	[b]		>2 cm	
12	Cutis		>1 cm	
13	Cutis	>2 cm		
14	Cutis + subc			>1 cm
15	Cutis	>2 cm[a]	+	

[a]Level of abnormality clinically estimated at the forearm, and at other locations indicated by +.
[b]Level of abnormality measured by vacuum method.

but there was no agreement on the other clinical signs being major or minor signs (data not shown).

DISCUSSION

Fifteen individuals (34%) out of the invited 41 took part as panelists in the first round. Only seven of them (47%) responded in the second round, and as a result, the planned Delphi process was not continued. Further, Cronbach's α for defining GJH, EDS-HT and JHS was 0.61, 0.79, and 0.44, respectively. The panelist to group correlation for the three clinical conditions, GJH, EDS-HT and JHS, ranged from −0.46 to 0.89, 0.03 to 0.68, and −0.72 to 0.68, respectively, indicating substantial heterogeneity among the panelists. The 5-point and the 9-point Beighton scoring systems were preferred by 2 or 13 panelists, respectively. Regardless of which scoring system was used, there was a considerable variation in opinion on how to perform the joint tests. The performance of other clinical tests essential for the diagnostic decision process, such as skin tests, also varied.

The Delphi method used for the study provides the opportunity to obtain participants opinion through anonymously answered questionnaires [Fink et al., 1984]. This method has a number of benefits: Panelists do not meet in person; no constraints on either size or composition of the panel; each panelist will have the same impact on the outcome of the group. However, the method also has some liabilities [Pill, 1971], for example, the IG can limit the scope of the issue evaluated by the panelists and thereby influence the consensus; as panelists never meet each other consensus can only be derived from information provided by the IG.

The response rate of 34%, was low compared to other studies [Graham et al., 2003; Dahmen et al., 2008; Dionne et al., 2008]. The invited panelists had recently performed or organized scientific studies in the hypermobility field and were expected to be familiar with and to have a genuine interest in solving the diagnostic problem. The questionnaire contained 50 main items, which were less than in other studies, namely 57 [Graham et al., 2003], respectively 77 items [Dionne et al., 2008]. Only one panelist complained about the method, commenting that it was difficult to give correct answers. The questionnaire contained also detailed questions about how to perform the clinical tests. It is possible that the questionnaire was too extensive, being one reason for the low response rate in the second round.

As described above the panelists came from a wide range of geographic locations, mostly in Europe, but also from Chile and USA, and covered all the medical specialties most commonly involved in the care of JHS and EDS. Such a panel composition ought to be optimal for the purposes of the study and with respect to gain acceptance from an international community of the obtained results [Fink et al., 1984]. However, this variation in panelists' specialization and nationality, could also be a reason for the low level of Cronbach's α and the broad variation in panelist to group correlation. Another reason could be the fact that Cronbach's α tends to be negatively influenced the lower the number of panelists is. Eventually, the low agreement could be due to the variability in the definitions of GJH and the overlap of criteria for BJHS and EDS-HT [Remvig et al., 2011].

The participants' different opinions about the performance of the Beighton tests could be due to the rather simple original description of the Beighton tests, that is, five photos with a short legend [Beighton and Horan, 1970], giving rise to individual interpretations of the approaches. Furthermore, it is worrying that two maneuvers with high reproducibility, that is, the 1st finger test (extended elbow and pronated hand) and the elbow test (abducted arm and supinated hand) [Juul-Kristensen et al., 2007], were less preferred. Also performance of tests for skin extensibility varied considerably, possibly due to varying descriptions of test performance and level for normal skin extensibility, ranging from 1.0 to 1.5 cm, tested on the volar aspect of the distal part of the forearm [Levy, 2004], to <4 cm at the volar aspect of the forearm [Holzberg et al., 1988]. Broad, atrophic scarring or papyraceous scarring is not one of the diagnostic criteria for EDS-HT [Beighton et al., 1998], but surprisingly, as many as 13 panelists recommended broad, atrophic scarring to have some importance for diagnosing EDS-HT, and seven (47%) even required papyraceous scarring in diagnosing EDS-HT. This could indicate confusion due to the extensive questionnaire or in worst-case poor knowledge of the criteria or lack of diagnostic expertise. In summary, both joint tests and skin tests are performed with varying maneuvers and with varying interpretations, and approximately 50% of the panelists (Table I) have not yet implemented tests shown to be reliable [Bulbena et al., 1992; Juul-Kristensen et al., 2007; Farmer et al., 2010; Remvig et al., 2010].

A high level of Cronbach's α indicates that the results of two independently selected panels from the same population will have a high correlation. Thus, a Cronbach's α >0.90 is not only desirable but also necessary when looking at items being part of syndrome criteria, that is, the diagnostic process [Bland and Altman, 1997]. A high level of Cronbach's α is also in accordance with other validated scales used in other clinical studies [Shrout and Fleiss, 1979].

CONCLUSION

The level of consensus for the importance of various items for diagnosing GJH, EDS-HT and JHS, among an international panel of clinicians with different education and different specialties, was well below the required limit for clinical decision-making and diagnosing. There was agreement on which tests to use, but performance of the tests (i.e., the specific maneuvers) varied considerably, and about 50% of the panelists still prefer maneuvers with unknown reliability, although reliable maneuvers are available. Consensus on tests and criteria through a Delphi process could not be obtained. There is an urgent need for better descriptions of the performance and interpretation of clinical tests used for these criteria sets, as well as more reliability studies on the various clinical maneuvers, tests and criteria commonly used. Subsequent intensive training and implementation of these tests and criteria, nationally as well as internationally should be established.

ACKNOWLEDGMENTS

We wish to acknowledge the focus group participants and the panelists (mentioned in alphabetical order) for taking time to make this study possible: Britta Berglund, Nurse; Eric Bjôrck, MD; Jaime Bravo, MD; Antonio Bulbena, MD; Marco Castori, MD; Richard Collinge, Physiotherapist; Raoul Engelbert, Physiotherapist; Mauro Ercolani, Psychologist; Stense Farholt, MD; Adam Farmer, MD; William Ferrell, MD; Rodney Grahame, MD; Elke S. Hjalmarsson, Physiotherapist; Maria Hoff, MD; Virginia B. Kraus, MD; Jens H. Kristensen, MD; Christina Kümmel, MD; L. Machet, MD; Anne L. Mitchell, MD; Jo Nijs, Physiotherapist; Svend Rand-Hendriksen, MD; Susanne Ulmann, MD; Lena L. Wekre, MD; and Pelin Yazgan, MD.

We also appreciate the advice given by Alan Hakim, MD, regarding the study protocol and the questionnaire.

REFERENCES

Beighton PH, Horan FT. 1970. Dominant inheritance in familial generalised articular hypermobility. J Bone Joint Surg 52:145–147.

Beighton P, de Pape A, Steinmann A, Tsipouras P, Wenstrup RJ. 1998. Ehlers–Danlos syndromes: Revised nosology. Villefranche 1997. Am J Med Genet 77:31–37.

Bland JM, Altman DG. 1997. Statistics notes: Cronbach's alpha. BMJ 314:572–573.

Bulbena A, Duro JC, Porta M, Faus S, Vallescar R, Martin-Santos R. 1992. Clinical assessment of hypermobility of joints: Assembling criteria. J Rheumatol 19:115–122.

Dahmen R, van der Wilden GJ, Lankhorst GJ, Boers M. 2008. Delphi process yielded consensus on terminology and research agenda for therapeutic footwear for neuropathic foot. J Clin Epidemiol 61:819–826.

Dalkey NC. 1969. The Delphi method: An experimental study of group opinion. Santa Monica: RAND Corporation.

Dalkey NC, Helmer O. 1963. An experimental application of the Delphi method to the use of experts. Manage Sci 9:458–467.

Dalkey NC, Brown B, Cochran S. 1969. The Delphi method III. Use of self-ratings to improve group estimates. Santa Monica: RAND Corporation.

Dionne CE, Dunn KM, Croft PR, Nachemson AL, Buchbinder R, Walker BF, Wyatt M, Cassidy JD, Rossignol M, Leboeuf-Yde C, Hartvigsen J, Leino-Arjas P, Latza U, Reis S, Gil Del Real MT, Kovacs FM, Oberg B, Cedraschi C, Bouter LM, Koes BW, Picavet HS, van Tulder MW, Burton K, Foster NE, Macfarlane GJ, Thomas E, Underwood M, Waddell G, Shekelle P, Volinn E, Von Korff M. 2008. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine 33:95–103.

Farmer AD, Douthwaite H, Gardiner S, Aziz Q, Grahame R. 2010. A novel in vivo skin extensibility test for joint hypermobility. J Rheumatol 37:1513–1518.

Fink A, Kosecoff J, Chassin M, Brook RH. 1984. Consensus methods: Characteristics and guidelines for use. Am J Public Health 74:979–983.

Graham B, Regehr G, Wright JG. 2003. Delphi as a method to establish consensus for diagnostic criteria. J Clin Epidemiol 56:1150–1156.

Grahame R, Bird HA, Child A. 2000. The revised (Brighton 1998) criteria for the diagnosis of benign joint hypermobility syndrome (BJHS). J Rheumatol 27:1777–1779.

Holzberg M, Hewan-Lowe K, Olansky A. 1988. The Ehlers–Danlos syndrome: Recognition, characterization, and importance of a milder variant of the classical form. J Am Acad Dermatol 19:656–666.

John MT. 2010. Improving TMD classification using the Delphi technique. J Oral Rehab 37:766–770.

Juul-Kristensen B, Røgind H, Jensen DV, Remvig L. 2007. Inter-examiner reproducibility of tests and criteria for generalised joint hypermobility and benign joint hypermobility syndrome. Rheumatology 46:1835–1841.

Levy H. 2004. Ehlers–Danlos syndrome, hypermobility type. GeneReviews. Available at: www.genetest.org.

Milholland AV, Wheeler SG, Heieck JJ. 1973. Medical assessment by a Delphi group opinion technique. N Engl J Med 288:1272–1275.

Pill J. 1971. The Delphi method: Substance, content, a critique and an annotated bibliography. Socio-Econ Plan Sci 5:57–71.

Remvig L, Jensen DV, Ward RC. 2007. Are diagnostic criteria for general joint hypermobility and benign joint hypermobility syndrome based on reproducible and valid tests? A review of the literature. J Rheumatol 34:798–803.

Remvig L, Duhn P, Ullman S, Arokoski J, Jurvelin J, Safi A, Jensen F, Farholt S, Hove H, Juul-Kristensen B. 2010. Skin signs in Ehlers–Danlos syndrome: Clinical tests and para-clinical methods. Scan J Rheumatol 39:511–517.

Remvig L, Engelbert RH, Berglund B, Bulbena A, Byers PH, Grahame R, Juul-Kristensen B, Lindgren KA, Uitto J, Wekre LL. 2011. Need for a consensus on the methods by which to measure joint mobility and the definition of norms for hypermobility that reflect age, gender and ethnic-dependent variation: Is revision of criteria for joint hypermobility syndrome and Ehlers–Danlos syndrome hypermobility type indicated? Rheumatology 50:1169–1171.

Shrout PE, Fleiss JL. 1979. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 86:420–428.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.