The Dyskinesia Impairment Scale, Second Edition: development, construct validity, and reliability

INTI VANMECHELEN[1]

ANNIKA DANIELSSON[2,3]

CECILIA LIDBECK[2,4]

KRISTINA TEDROFF[2]

ELEGAST MONBALIU[1]

LENA KRUMLINDE-SUNDHOLM[2]

1 Department of Rehabilitation Sciences, KU Leuven, Bruges, Belgium; **2** Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden; **3** Sachs' Children and Youth Hospital, Stockholm South General Hospital, Stockholm, Sweden; **4** Neuropediatric Department, Karolinska University Hospital, Stockholm, Sweden.

Correspondence to Inti Vanmechelen, Department of Rehabilitation Sciences, KU Leuven, Bruges Campus, Spoorwegstraat 12, 8200 Bruges, Belgium. E-mail: inti.vanmechelen@kuleuven.be

ABBREVIATIONS

DIS	Dyskinesia Impairment Scale
DIS-II	Dyskinesia Impairment Scale, Second Edition
DIS-II-CA	Dyskinesia Impairment Scale, Second Edition-choreoathetosis subscale
DIS-II-D	Dyskinesia Impairment Scale, Second Edition-dystonia subscale
PCA	principal component analysis

[abstract]

AIM To create a shortened, more user-friendly Second Edition of the Dyskinesia Impairment Scale (DIS-II) to assess dystonia and choreoathetosis and evaluate its construct validity and reliability.

METHOD Scale development included an online expert meeting ($n = 21$) and iterative discussions within the research group ($n = 6$). A Rasch measurement model analysis on DIS scores from individuals with dyskinetic cerebral palsy or inherited/idiopathic dystonia ($n = 123$, 74 males, mean age = 14 years, SD = 5 years) was performed to evaluate the construct validity and reliability of the DIS-II.

RESULTS The DIS-II evaluates dystonia and choreoathetosis in action and rest in 11 body regions, with action items scored from 0 to 3 and rest items 0 to 2. The number of videos to record are reduced from 26 to 14 and the items to score are reduced from 144 to 88. Rating scale functioning, goodness-of-fit evaluation, principal component analysis, and targeting met the predefined quality criteria of the study and construct validity was therefore considered good. Furthermore, person reliability indicated that the DIS-II can separate individuals into eight distinct ability levels.

INTERPRETATION The DIS-II provides valid and reliable measures for dystonia and choreoathetosis and reduces the administration and scoring time compared with the DIS. The DIS-II logit scores (interval level data) enhance comparison over time and between individuals in clinical practice and research.

What this paper adds

- Compared with the Dyskinesia Impairment Scale (DIS), the shortened Second Edition (DIS-II) requires half of the videos to be recorded.
- The DIS-II has a simplified rating scale, requiring scoring of 88 instead of 144 items.
- The DIS-II has shown excellent reliability and good construct validity.
- The interval properties of the DIS-II are superior to the ordinal level outcome measures of the DIS.

[main]

Dystonia is a movement disorder in which involuntary sustained or intermittent muscle contractions cause twisting and repetitive movements, abnormal postures, or both.[1] Choreoathetosis can be separated into chorea and athetosis, where chorea is defined as rapid involuntary, jerky, and fragmented movements, while athetosis consists of slower, constantly changing, writhing, or contorting movements.[1] Dystonia and choreoathetosis are prevalent in dyskinetic cerebral palsy (CP) and in different inherited or acquired conditions.[2,3]

Assessment of dystonia and choreoathetosis through a reliable and valid scale is fundamental for research and in clinical decision-making. Assessment scales for dystonia are available for different populations, for example, for dystonia in children with CP or for inherited or idiopathic dystonia in adults,[4–6] but their psychometric properties have been questioned.[7,8] Only one scale that evaluates both dystonia and choreoathetosis is available—the Dyskinesia Impairment Scale (DIS)—which was designed for children and young adults with dyskinetic CP.[9]

The DIS is compatible with the consensus CP classification and definitions of dystonia and choreoathetosis.[1,10,11] It has shown good-to-excellent psychometric properties, suggesting that it is a promising tool in intervention studies and when objectively evaluating disease

progression.[9,12,13] Recently, the DIS was also shown to be valid and reliable in children and young adults with inherited or idiopathic dystonia.[14] However, the DIS is time-consuming to complete and therefore often disregarded in clinical practice. Furthermore, the result is presented as a percentage score, making comparison over time or between individuals complicated.

The primary objective of this study was to develop a shortened, user-friendly Second Edition of the DIS, the DIS-II, and evaluate its construct validity and reliability. The secondary objective was to adjust the percentage outcome measure in the original DIS to an absolute outcome measure in the DIS-II to enable comparison over time and between individuals.

METHOD

Instruments

The original DIS consists of two subscales, one for dystonia and one for choreoathetosis, which evaluate the presence of dystonia/choreoathetosis in 12 body regions during two actions and one rest posture (Table 1).[9] The DIS is scored from 26 video recordings of 30 seconds each according to a standardized protocol. The amount of time (duration) and the range of motion (amplitude) of dystonia/choreoathetosis are assessed on a 5-step ordinal scale (0–4).[9] Zero indicates no dystonia/choreoathetosis and 4 indicates dystonia/choreoathetosis always present or in a maximal range of motion (>90%). The total DIS score is the sum of the dystonia and choreoathetosis subscales and ranges from 0 to 576 points. Depending on the ability level of the individual being evaluated, only the actions/rest postures they can perform are assessed. Hence, the result of the DIS is presented as a percentage score for DIS and the dystonia and choreoathetosis subscales.

Table 1: Actions and rest postures in the original DIS versus the DIS-II

Body region	DIS		DIS-II	
	Action	**Rest**	**Action**	**Rest**
Eyes	**Eye tracking** Eye blinking	Eyes in sitting	Eye blinking	Eyes in sitting
Mouth	**Mouth open/close** Speech	Mouth in sitting	Speech	Mouth in sitting
Neck	**Lateroflexion** Rotation	Neck in sitting	Rotation	Neck in sitting
Trunk	**Active sitting** **Forward flexion**	**Trunk in sitting**	– –	– –
Arm proximal	**Arm abduction** Grasp/reach a pen (supine)	Arm proximal in sitting	Grasp/reach a pen (supine)	Arm proximal in sitting
Arm distal	Grasp and move a cup **Grasp and move a pen**	Arm distal in sitting	Grasp and move a cup	Arm distal in sitting
Leg proximal	Rolling **Standing**	Leg proximal in lying	Rolling	Legs proximal in lying
Leg distal	**Rolling** Heel/toe tapping	Leg distal in lying	Heel/toe tapping	Legs distal in lying

Bold text signifies deleted items. Abbreviations: DIS, Dyskinesia Impairment Scale; DIS-II, Dyskinesia Impairment Scale, Second Edition.

Procedure

This study consisted of three phases: (1) development of the DIS-II; (2) data collection; and (3) Rasch analysis of the DIS-II.

Development of the DIS-II

Clinicians and researchers using the DIS were invited to participate in an advisory expert meeting. The invited experts were contacted via e-mail through the European Dyskinetic Cerebral Palsy Network, a network formed to improve research and clinical implementation in

dyskinetic CP via international collaborations. This virtual whole-day meeting gathered 21 experts from five countries in February 2021. The attendees individually answered questions concerning the number of body regions and actions/rest postures, action selection, and rating scale structure. The questions were followed by an open discussion covering the aforementioned aspects of the DIS. The results from this meeting were analysed and discussed within the research group that includes the developer of the original DIS (EM). Consensus was reached through iterative discussions within the research group, resulting in a proposal for the DIS-II for later evaluation with Rasch analysis.

Data collection for the Rasch analysis

DIS scores from 123 individuals with dyskinetic CP, inherited/idiopathic, or acquired dystonia were collected from Belgium (n = 100) and Sweden (n = 23) with ages ranging from 5 to 25 years. For 97 out of the 123 participants, scores from earlier studies could be reused. The remaining 26 participants had been recorded according to the DIS protocol for clinical purposes without being scored and these videos were scored for this study. For the Rasch measurement model analysis, a sample size of 100 persons should indicate item calibration within plus-minus 0.5 logit with 95% confidence.[15] Participant characteristics are given in Table 2.

Table 2: Participant characteristics

Characteristic	**Participants (n = 123)**
Sex, *n*	
Male	74
Female	49
Age, mean (SD), range, years	14 (5), 5–25
Diagnosis, *n*	
Dyskinetic cerebral palsy	99
Inherited/idiopathic dystonia	22

Acquired dystonia	2
GMFCS level (*n*)	
I	21
II	17
III	13
IV	28
V	44
MACS level (*n*)	
I	18
II	18
III	19
IV	22
V	44
Unknown	2

Abbreviations: GMFCS, Gross Motor Function Classification System; MACS, Manual Ability Classification System.

Rasch analysis of the DIS-II

A Rasch measurement model analysis was used to investigate the DIS-II, using WINSTEPS v4.8.2.0 (https://www.winsteps.com/winsteps.htm). Construct validity, including rating scale functioning, principal component analysis (PCA), fit statistics, point-measure correlations, and targeting as well as aspects of reliability were investigated using the Rasch model and ordinal DIS scores were converted to interval measures in the unit logits. The grouped Rating Scale Model was used for the analysis because all DIS action items are assessed with the same rating scale, with one category ('action not done') less for the DIS rest items.[16,17] In the DIS-II, a distinction is made between the left and right extremities; however, for the purpose of item hierarchy, left and right were converted to 'best' and 'worst', where 'best' was the item with

the lowest score for dystonia or choreoathetosis. First, the total DIS-II scale was analysed; subsequently, the dystonia (DIS-II-D) and choreoathetosis **(**DIS-II-CA) subscales were analysed separately.

The latent trait in this assessment is the amount of dystonia/choreoathetosis. A high person measure indicates more dystonia/choreoathetosis than a low measure. A high item measure indicates less dystonia/choreoathetosis than items with a low measure. Test items are the duration/amplitude of dystonia/choreoathetosis during action/rest in a specific body region.

First, rating scale functioning was investigated. The quality criteria used for an appropriate rating scale were: (1) average measure advancing monotonically by category; (2) item step calibration increasing with higher categories; (3) each category being represented by a curve with a distinct peak on visual inspection; (4) frequency for all four categories exceeding 10 ratings on average; and (5) the outfit mean square should not exceed 2.0 for any category.[18–20]

Second, unidimensionality was investigated through PCA of the residuals.[21] The variance in raw scores explained by the amount of dystonia and choreoathetosis represents the first dimension, which is our dimension of interest. PCA of the residuals is performed to identify a possible second or third dimension in the remaining unexplained variance. A high unexplained variance in the first contrast indicates a second dimension, and should be further analysed. Following quality criteria were applied, 3% to 5% unexplained variance in the first contrast was considered very good, 5% to 10% good, and 10% to 15% fair.[21] Goodness-of-fit to the Rasch model was examined by fit statistics and point-measure correlations. Fit statistics indicate how accurately or predictably data fit the model, that is, misfit indicates that the model does not measure the latent variable sufficiently. Misfit can be explained as outfit or infit misfit, where the outfit statistic is an unweighted estimate of the degree of misfit, whereas the infit statistic is weighted to give more value to observations near the target/measure of the item/person. Since misfitting infit statistics are a greater threat to test validity than misfitting

outfit statistics, we considered item and person misfit as infit mean-square residual values equal to or greater than 1.4 in combination with standardized z values equal to or greater than 2.0.[22] A 95% fit was considered evidence of goodness-of-fit of items and persons to the Rasch model. Item point-measure correlations were used to investigate how well item measures represent the latent trait; positive point-measure correlation between items are considered sufficient.[23]

Third, targeting between persons and item measures was investigated through visual inspection of the item-person map to see whether item difficulty calibrations were spread evenly along the whole measurement range covering all levels of person ability. Ceiling and floor effects were investigated by checking if an individual scored maximum or minimum on the test. Additionally, a difference between person and item means that a difficulty of less than 0.5 logits is considered a quality criterion for good targeting, where the mean of the item measured is 0 by definition.[24]

Fourth, the test reliability of item and person measures was determined by evaluating internal consistency (corresponding to Cronbach's alpha) and whether the items separate persons into distinct levels of ability (i.e. amount of dystonia/choreoathetosis in this study). Person reliability is interpreted as the ability of the scale to reliably rank individuals relative to the location of the measure. The item reliability coefficient is a measure of internal consistency, providing the degree of confidence that could be placed in the consistency of the estimates: values greater than 0.80 were considered good, greater than 0.90 excellent.[2,25] Separation was calculated to investigate how many statistically distinguishable measurement levels existed in the sample; it should be greater than 2 to separate people into at least two ability levels. Person strata measures were calculated using following formula:[26]

Strata = 4*(separation index + 1)/3

Person strata of 4 to 5 was regarded as very good and greater than 5 as excellent.[26]

Ethics

This study obtained ethical approval from the research ethics committee of UZ/KU Leuven (no. S64764), the regional ethical review board in Stockholm (day-book no. 2016/682-31), and the Swedish ethical review authority (day-book no. 2020-05271), with written informed consent for the research participants and oral informed consent for the clinical patients which was in line with requirements of the ethical committee.

RESULTS

Development of the DIS-II

Expert meeting

The attendees of the expert meeting (n=21) were physiotherapists (n=11), occupational therapists (n=2), and paediatric neurologists (n=8). All had experience with the DIS: in research (n=4), in the clinic setting (n=4), and both in research and clinical setting (n=12) (Twenty votes instead of 21). Nineteen of the attendees had more than 5 years of experience in evaluating dystonia and choreoathetosis, twelve had more than 10 years of experience.

Reduction of body regions and actions

Sixty-nine per cent of the attendees at the expert meeting wanted to reduce the number of body regions to be scored in the DIS but there was no obvious agreement as to which body region should be removed. Based on the high number of individuals unable to perform trunk actions and the previously reported low interrater reliability values,[12,14] the research team decided to delete the trunk region. Eighty-seven per cent of the attendees were in favour of deleting one of the two actions to be scored. The research team subsequently analysed each action from a clinical point of view, additionally taking into account scoring feasibility. As a result, reduction of body regions from 12 to 11 and reduction of one action per body region, reduced the total number of videos to be recorded from 26 to 14. Additionally, the total number of test items was reduced from 144 in the original DIS to 88 in the DIS-II (Table 1).

Reduction of rating scale

Ninety-three per cent of the attendees agreed that the number of scale steps should be reduced. In the original DIS, an action item is scored only if the person can perform the action, otherwise the item receives '/'. In the DIS-II, a scoring step for items 'not being done' was added. This served two purposes. First, percentage scores are not needed in the DIS-II making comparison easier. Second, conversion from raw scores to logit measures is only possible when all items are scored. To convert raw scores to Rasch logits by using a transformation table, all items need to be scored. Hence, a scoring step for items not being done was added in the DIS-II. Two different rating scale options were identified as clinically relevant by the research group and both options were analysed with the Rasch measurement model analysis to choose the most appropriate one for the DIS-II (Table 3). The DIS-II score form and video protocol can be found in Appendix S1.

Table 3: Scoring scale steps in the original DIS compared with the two options for the DIS-II

DIS = Dyskinesia Impairment Scale. DIS-II = Dyskinesia Impairment Scale, Second Edition

Original DIS	**DIS-II option 1**	**DIS-II option 2**
0 = no dystonia/choreoathetosis	0 = no dystonia/choreoathetosis	0 = no dystonia/choreoathetosis
1 = <10%	1 = <50%	1 = <50%
2 = ≥10% to <50%	2 = ≥50% to <90%	2 = ≥50%
3 = ≥50% to <90%	3 = ≥90%	3 = not done
4 = ≥90%	4 = not done	
/ = not done		

Rasch measurement model analysis

Rasch summary statistics for the total DIS-II scale were similar for both rating scale options with person reliability being 0.97 and person strata being 8, whereas item reliability was 0.96 for option 2 and 0.97 for option 1. A similar trend was observed for the DIS-II-D and DIS-II-

CA. However, rating scale functioning showed obvious differences between the two rating scale step options (Table S1). For the DIS-II, visual inspection of a distinct peak for each category was obvious for option 2 but not for option 1, where there was no peak for category 2 (Figure S1). Furthermore, the number of disordered thresholds when analysing each test item separately was much higher in option 1 versus 2 (16 vs 2). All further results are thus based on option 2, reducing the total score from 576 raw scores (original DIS) to 220 in the DIS-II. From a total of 10 824 scored items, 12 values were missing.

Rating scale functioning

For the DIS-II scale and both subscales, the average measure advanced by category and item step calibration increased with higher categories for all items. There were no average counts lower than 10 per category in any of the scales and distinction of peaks for each category was confirmed for all scales on visual inspection. Finally, the outfit mean square for each category was lower than 2 for all scales.

Evaluation of unidimensionality for the DIS-II showed 45.1% variance explained by the measures and 9.0% unexplained variance in the first contrast for the total scale. For the DIS-II-D and DIS-II-CA subscales, the variance explained by measure was 49.6% and 41.1% respectively with 8.1% and 12.6% variance in the first contrast. Since a first contrast higher than 5% may be associated with multidimensionality, the principal component factor plot was analysed to explore multiple dimensions in the data. In the DIS-II and DIS-II-D/CA, a second dimension was clearly present, that being action and rest. Additionally, in the PCA of the DIS-II there were indications of a third dimension, that being dystonia versus choreoathetosis.

Model fit

For both the DIS-II and the DIS-II-CA, the items ‘eyes: choreoathetosis action duration’ and ‘eyes: choreoathetosis action amplitude’ were misfitting. For the DIS-II-D, ‘worst leg distal rest duration’ was misfitting (Table S2). On a total of 88 items for the DIS-II and 44 items for

the DIS-II-D/CA, this implies more than 95% item fit for all scales (Tables 3 and S3). All item point-measure correlations were positive and ranged from 0.16 to 0.75. For person misfit, there were 10, 14, and 15 misfitting persons for the DIS-II, DIS-II-D, and DIS-II-CA respectively. Since this is higher than 5% of the total sample, the analysis was redone after removing the misfitting persons. This did not alter the summary statistics; additionally, new misfitting persons appeared. Thus, all persons were kept for all further analyses since misfitting persons did not have an impact on the results. Regarding targeting, the person-item bar chart showed a good spread of item difficulties covering all person abilities; floor and ceiling effects were absent (Figure S2). The person measure mean was −0.26 for the DIS-II, and 0.31 and −0.82 for the DIS-II-D and DIS-II-CA, which indicates good targeting for the DIS-II and the DIS-II-D.

Test reliability

Internal consistency values were excellent for both items and persons for the DIS-II and the DIS-II-D/CA, as well as item separation index values (Table 4). Person separation strata were excellent for the DIS-II and very good for the DIS-II-D/CA.

Table 4: Rasch analysis results for the total DIS-II, DIS-II-D, and DIS-II-CA

	DIS-II	**DIS-II-D**	**DIS-II-CA**
Principal component analysis			
Variance explained by measures (%)	45.1	49.6	41.1
Unexplained variance in the first contrast (%)	9.0	8.1	12.6
Person			
Measure, mean (SD)	−0.26 (1.06)	0.31 (1.46)	−0.82 (1.18)
Measure range	−4.13 to 2.41	−4.24 to 4.72	−6.10 to 1.70
Mean standard error	0.16	0.26	0.25
Reliability	0.97	0.96	0.93
Separation index	5.77	4.63	3.6
Strata	8.03	6.5	5.13
Items			
Measure, mean (SD)	0 (0.73)	0 (0.56)	0 (0.53)
Measure range	−1.76 to 1.36	−1.31 to 1.24	−0.97 to 1.11
Mean standard error	0.13	0.15	0.14
Reliability	0.96	0.92	0.92
Separation index	5.06	3.42	3.46

Abbreviations: DIS-II, Dyskinesia Impairment Scale, Second Edition; DIS-II-D, Dyskinesia Impairment Scale, Second Edition-dystonia subscale; DIS-II-CA, Dyskinesia Impairment Scale, Second Edition-choreoathetosis subscale.

DISCUSSION

In this study, where the primary objective was to shorten the original DIS scale, the more user-friendly and shorter DIS-II was developed. The number of videos to record and score were reduced by approximately 50% and the rating scale was simplified. The DIS-II evaluates dystonia and choreoathetosis with excellent item and person reliability and separates individuals into eight distinct ability levels. Rating scale functioning, PCA, fit statistics, and targeting met the predefined quality criteria of the study and therefore provided support for construct validity of the DIS-II. Additionally, DIS-II logit scores enhanced comparison of the presence of dystonia and choreoathetosis over time or between individuals in the clinic and in research.

For the patient, the reduction in number of videos to be recorded in the DIS-II is beneficial, allowing more widespread use of the scale due to its lower time and effort burden. For clinicians, the long scoring time of the DIS currently hampers its implementation in daily practice. Although the DIS provides much useful information, clinicians often cannot exploit its full potential due to time restrictions. The DIS-II provides a shorter, more user-friendly alternative with reduced scoring times owing to a simplified rating scale structure and a reduction of body regions and actions. The time to score the DIS-II is approximately 45 to 60 minutes. The reduction of actions and body regions decreases the items to be scored, whereas the simplified rating scale structure is expected to decrease scoring time for each video sequence because the choice of response options is lower.

Rasch analysis of the DIS-II showed good item fit with only two misfitting items: namely amplitude and duration of choreoathetosis in the eye region during action. Thus, one could argue that those two items could be removed; however, we chose not to do so because symmetry between the DIS-II-D and DIS-II-CA was regarded highly important by the expert group and a prerequisite of the scale. Hence, if items assessing choreoathetosis in the eye region

during action were removed, then items assessing dystonia during action would also be removed automatically. Dystonia affects quality of life and participation to a greater extent than choreoathetosis,[27] supporting our choice of keeping the possibility to assess dystonia in the eye region. However, when interpreting choreoathetosis in the eye region, we encourage caution. Regarding the evaluation of dystonia/choreoathetosis in the trunk region, we acknowledge the importance of this body region but we decided to remove the trunk region in the DIS-II because a high number of participants in our data set could not perform trunk actions; therefore, this body region distorted the scale construct.

The unidimensionality of testing means that differences in observed scores only depend on the latent trait, in this case dystonia and choreoathetosis. PCA of the latent trait was used to evaluate this aspect. In this study, the unexplained variance in the first contrast indicated a possible second dimension, that is, action and rest. This dimension was expected since action and rest are two important aspects of dystonia and choreoathetosis. Thus, the presence of this second dimension does not violate the unidimensionality of the scales. For the DIS-II, dystonia and choreoathetosis appeared as the third dimension, which agrees with the clinical separation, distinct definitions, and development of subscales for these movement disorders and is subsequently not a threat to the unidimensionality of our scale.

The secondary objective of developing the DIS-II was to adjust the percentage outcome measure to an absolute outcome measure. In the original DIS, the result of an assessment is presented as raw or percentage scores, excluding actions/rest postures that were ‘not done’. Both scores are dependent on the number of actions performed; thus, they cannot be compared between individuals or within one individual over time. However, the new rating scale in the DIS-II, including a score for each item where actions/rest postures were not done, leads to an outcome measure that enables an unbiased comparison independent of the number of actions included in the assessment. The results can be presented as raw scores giving ordinal data or

logit measures by using conversion tables (Table S4). Logit measures are superior to raw scores because they are interval data whereas raw scores are ordinal data. In the clinic, the DIS-II can thus be used to follow disease course over time or evaluate medical interventions. Presenting the results as raw scores is still possible if the assessor wants to know the presence of dystonia/choreoathetosis during rest or action separately.

The current study has some limitations. First, our sample size did not allow either differential item functioning analysis or differential person functioning analysis because a sample size of 100 persons in each group is required.[28] However, the group of individuals with dystonia and choreoathetosis is limited, which affected the sample size included in this study. Notwithstanding, current study holds the biggest sample in children and adolescents with dystonia and choreoathetosis. Second, the DIS-II has not been assessed for intrarater, interrater, and test–retest reliability and further studies are warranted to assess these research questions. However, a strength of this study is the use of an advisory expert meeting and Rasch analysis to develop the DIS-II. Before the DIS-II can be considered a truly robust tool, the use of the new 3/4-point rating scale should be evaluated. The current results are built on merging categories, not from direct observation. Additionally, future research should focus on responsiveness to change, for example, before and after treatment with intrathecal baclofen or deep brain stimulation and on concurrent validity, for example, relationship to other measures of dystonia and choreoathetosis.

Conclusion

The shortened, user-friendly DIS-II reduces both administration and scoring time, thereby increasing its applicability in both clinical setting and research. Rasch analysis provides support for construct validity of the DIS-II and the excellent person reliability shows that the scale can be used to separate the amount of dystonia and choreoathetosis into eight distinct levels and thereby has the potential to capture change.

ACKNOWLEDGEMENTS

We thank the attendees of the DIS advisory expert meeting for their time and effort in discussing aspects of the DIS.

This study was funded by an aspirant fellowship of the Research Foundation-Flanders, by Sällskapet Barnavård, and by the Sachs' Children and Youth hospital.

Supporting information

The following additional material may be found online:

Table S1: Category average function for option 1 and 2 of the DIS-II.

Table S2: Total DIS-II item statistics in measure order.

Table S3: Item fit statistics in measure order for the DIS-II-CA and DIS-II-D.

Table S4: Conversion table for the total DIS-II.

Figure S1: Category peaks for DIS-II option 2 rest, DIS-II option 1 action, DIS-II option 2 action, and DIS-II option 1 rest.

Figure S2: Item targeting for total DIS-II, DIS-II-D, and DIS-II-CA.

Appendix S1: DIS-II score form.

REFERENCES

1. Sanger TD, Chen D, Fehlings DL, Hallett M, Lang AE, Mink JW, et al. Definition and classification of hyperkinetic movements in childhood. Mov Disord. 2010;25(11):1538-49.

2. Monbaliu E, Himmelmann K, Lin J-P, Ortibus E, Bonouvrié L, Feys H, et al. Clinical presentation and management of dyskinetic cerebral palsy. The Lancet Neurology. 2017;16(9):741-9.
3. Albanese A, Bhatia K, Bressman SB, Delong MR, Fahn S, Fung VS, et al. Phenomenology and classification of dystonia: a consensus update. Mov Disord. 2013;28(7):863-73.
4. Burke RE, Fahn S, Marsden CD, Bressman SB, Moskowitz C, Friedman J. Validity and reliability of a rating scale for the primary torsion dystonias. Neurology. 1985;35(1):73-7.
5. Comella CL, Leurgans S, Wuu J, Stebbins GT, Chmura T. Rating scales for dystonia: a multicenter assessment. Mov Disord. 2003;18(3):303-12.
6. Barry MJ, VanSwearingen JM, Albright AL. Reliability and responsiveness of the Barry-Albright Dystonia Scale. Dev Med Child Neurol. 1999;41(6):404-11.
7. Stewart K, Harvey A, Johnston LM. A systematic review of scales to measure dystonia and choreoathetosis in children with dyskinetic cerebral palsy. Dev Med Child Neurol. 2017;59(8):786-95.
8. Albanese A, Sorbo FD, Comella C, Jinnah HA, Mink JW, Post B, et al. Dystonia rating scales: critique and recommendations. Mov Disord. 2013;28(7):874-83.
9. Monbaliu E, Ortibus ELS, de Cat JOS, Dan B, Heyrman L, Prinzie P, et al. The Dyskinesia Impairment Scale: a new instrument to measure dystonia and choreoathetosis in dyskinetic cerebral palsy. Developmental Medicine & Child Neurology. 2012;54(3):278-83.
10. Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, et al. A report: the definition and classification of cerebral palsy April 2006. Dev Med Child Neurol Suppl. 2007;109:8-14.

11. Cans C, Dolk H, Platt MJ, Colver A, Prasauskiene A, Krageloh-Mann I, et al. Recommendations from the SCPE collaborative group for defining and classifying cerebral palsy. Dev Med Child Neurol. 2007;49:35-8.

12. Monbaliu E, Ortibus E, Prinzie P, Dan B, De Cat J, De Cock P, et al. Can the Dyskinesia Impairment Scale be used by inexperienced raters? A reliability study. Eur J Paediatr Neurol. 2013;17(3):238-47.

13. Vanmechelen I, Dan B, Feys H, Monbaliu E. Test-retest reliability of the Dyskinesia Impairment Scale: measuring dystonia and choreoathetosis in dyskinetic cerebral palsy. Dev Med Child Neurol. 2020;62(4):489-93.

14. Danielsson A, Vanmechelen I, Lidbeck C, Krumlinde-Sundholm L, Ortibus E, Monbaliu E, et al. Reliability and Validity of the Dyskinesia Impairment Scale in Children and Young Adults with Inherited or Idiopathic Dystonia. J Clin Med. 2020;9(8).

15. Linacre J. Sample size and item calibration stability. Rasch Measurement Transactions Contents. 1994;7:1.

16. Kottorp A, Petersson I. Psychometric evaluation of an assessment of awareness using two different Rasch models. Scand J Occup Ther. 2011;18(3):219-30.

17. Linacre JM. A user-guide to WINSTEPS: Grouped rating-scale model. Chicago, IL; 2009 [Available from: https://www.winsteps.com/winman/isgroups.htm].

18. Linacre JM. Optimizing rating scale category effectiveness. J Appl Meas. 2002;3(1):85-106.

19. Bond T, Fox C. Applying the Rasch model: Fundamental Measurment in the Human Sciences. Second edition ed. New Yersey: Lawrence Erlbaum Associates 2007. 340p.; p.222-224.

20. Fisher W. Rating scale instrument quality criteria. Rasch Measurement Transactions Contents. 2007;21:1.

21. Bond T, Fox C. Applying the Rasch model: Fundamental Measurment in the Human Sciences, Second edition ed. New Yersey: Lawrence Erlbaum Associates 2007. 340p.; p.251-260.

22. Bond T, Fox C. Applying the Rasch model: Fundamental Measurement in the Human Sciences. New Yersey: Lawrence Erlbaum Associates 2007. 340p.; p.238-239.

23. Linacre J. A users guide to WINSTEPS. Rasch-model computer programs: Point-measure correlations. Chicago, IL; 2009 [Available from: https://www.winsteps.com/winman/correlations.htm].

24. Fisher Jr. WP. Rating Scale Instrument Quality Criteria. Rasch Measurement Transaction 2007 [Available from: http://www.rasch.org/rmt/rmt211a.htmhttp://www.rasch.org/rmt/rmt211a.htm].

25. Bond T, Fox C. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. New Jersey: Lawrence Erlbaum Associates; 2007. 340p.; p.40-41.

26. Linacre J. Number of person or item strata. Rasch Measurement Transactions Contents. 2002;16:888.

27. Monbaliu E, de Cock P, Ortibus E, Heyrman L, Klingels K, Feys H. Clinical patterns of dystonia and choreoathetosis in participants with dyskinetic cerebral palsy. Dev Med Child Neurol. 2016;58(2):138-44.

28. Linacre JM. Differential Item Functioning DIF Sample Size Nomogram. 2013 [Available from: https://www.rasch.org/rmt/rmt264a.htm