



RESEARCH REPORT

Rasch-built overall disability scale for POEMS syndrome (POEMS-RODS)

Ryan Yann Shern Keh^{1,2}  | James B. Lilleker^{2,3} | Tim Lavin² | Jonathan Sive⁴ | Shirley D'Sa⁴ | Gita Ramdharry^{1,5} | Aisling S. Carr^{1,5} | Michael P. Lunn^{1,5} 

¹Queen Square Centre for Neuromuscular Diseases, National Hospital of Neurology and Neurosurgery, Queen Square, University College London Hospitals NHS Foundation Trust, London, UK

²Manchester Centre for Clinical Neurosciences, Salford Royal Hospital, Northern Care Alliance NHS Foundation Trust, Manchester, UK

³Division of Musculoskeletal and Dermatological Sciences, Manchester Academic Health Sciences Centre, University of Manchester, Manchester, UK

⁴Department of Haematology, University College London Hospitals NHS Foundation Trust, London, UK

⁵Institute of Neurology, University College London, London, UK

Correspondence

Michael P. Lunn, Queen Square Centre for Neuromuscular Diseases, National Hospital of Neurology and Neurosurgery, Queen Square, University College London Hospitals NHS Foundation Trust, London, UK.
Email: michaellunn@nhs.net

Abstract

Patient-reported outcome measures engage patients in disease severity measurement and the metrics reported can be meaningful to their lives. The Polyneuropathy, Organomegaly, Endocrinopathy, Monoclonal protein and Skin changes syndrome (POEMS) is a complex multisystem disorder with disabling neuropathy which is distinct from other acquired inflammatory neuropathies. No current POEMS-specific validated disability scales exist. To address this, we have produced a Rasch-built overall disability scale (RODS) specific to POEMS. A 146-item preliminary questionnaire containing relevant activity and participation items for neuropathic disability was applied to 49 clinically stable patients with POEMS from the UK national POEMS cohort. A total of 123 items not fulfilling Rasch model expectations were sequentially removed. The final 23-item POEMS-RODS fulfilled Rasch model expectations and showed acceptable test-retest reliability. The 23-item POEMS-RODS is a disease-specific patient-reported outcome measure able to detect activity limitations within the range of ability demonstrated by the UK POEMS cohort. Larger international studies are needed to confirm the broader applicability and responsiveness of this scale in other countries.

KEYWORDS

neuropathy, POEMS syndrome, Rasch-built overall disability scale

1 | INTRODUCTION

The Polyneuropathy, Organomegaly, Endocrinopathy, Monoclonal protein and Skin changes syndrome (POEMS) is a rare, complex multi-system disease, often first diagnosed by neurologists because of the disabling neuropathy which coexists with a monoclonal plasma cell disorder.¹ Although sometimes misdiagnosed as chronic inflammatory demyelinating polyradiculoneuropathy (CIDP), the neuropathic distribution is distinct from CIDP, and the treatment, pattern of disability and outcomes are different. Serum or plasma Vascular Endothelial Growth Factor has good diagnostic sensitivity and specificity for

POEMS and is sometimes useful for relapse identification, but it does not reflect the disability or severity of the disease.²

The quantitative and meaningful measurement of disease severity assists clinical diagnosis, prognosis and assessment of natural history and response to intervention. Classical clinical test theory has generated hundreds of scales that may be useful but are often non-linear, statistically invalid for populations, not relevant to patient outcomes or unresponsive to change. Modern clinimetrics, for example, using item response theories such as Rasch measurement theory,³ has evolved to support the development of many improved scales. The Rasch method uses responses to ordinal (ordered) questions from a

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Journal of the Peripheral Nervous System* published by Wiley Periodicals LLC on behalf of Peripheral Nerve Society.

selected cohort of individuals to build a single-dimension, linear scale of item difficulty and person ability.

The most meaningful health domain for patients is ability/disability. Patient-reported outcome measures (PROMS) engage the patient in the process of disease measurement. Rasch-built Overall Disability Scales (RODS) for neuropathies were introduced by the PeriNOMS group for GBS, CIDP, multifocal motor neuropathy (MMN) and para-proteinemic neuropathies (the inflammatory neuropathy RODS, or i-RODS, and MMN-RODS)^{4,5} and have demonstrated their relevance, responsiveness and statistical validity over other, non-linear disease scores. However, these scales are validated for specific diseases and thus one scale is not valid for alternative conditions. No POEMS-specific validated disability scales exist. This makes it challenging to measure disease severity in the clinical setting, and to demonstrate treatment effectiveness in research trials.

We have built a RODS specific to POEMS neuropathy (POEMS-RODS), using the UK national POEMS cohort, which conforms to the Rasch unidimensional model and reflects the range of neuropathic disabilities caused by POEMS.

2 | MATERIALS AND METHODS

The University College London Hospitals multidisciplinary team runs the national POEMS clinic, which offers specialist assessment and follow-up to UK patients with POEMS. Over 130 patients have been seen since 2000, with approximately 10 to 15 new diagnoses yearly and about 70 patients under ongoing follow-up.

We administered 146-question preliminary RODS, to all responders within the cohort. This item bank was developed by the PeriNOMS group by reviewing activity and participation items in the International Classification of Functioning, Disability and Health, and selecting those items expected to be relevant to patients with neuropathy.⁴ It has been successfully used in the creation of other Rasch-built neuropathy disability scales including the i-RODS and the MMN-RODS.^{4,5} As POEMS-related disability is primarily neuropathic,⁶ the complete item bank most likely contains items reflecting the range of disability demonstrated by individuals with POEMS. Participants completed this as an online questionnaire, on a paper form (in a clinic or returned by post), or over the telephone. Participants were asked to respond whether they found the tasks 'possible, without any difficulty [2]', 'possible, but with some difficulty [1]', 'not possible to perform [0]', or 'not applicable [9]'. Participants were asked to respond on their ability to perform the tasks with any walking aids and/or orthotics that they usually used in daily activities.

Age and gender of participants were recorded as person factors, arbitrarily categorised as age (<50 years, 50-69 years, >70 years) and gender (male vs female) to enable an analysis of differential item functioning in the Rasch model.

2.1 | Rasch analysis

In the initial preliminary RODS, items scored as 'not applicable [9]' were interpreted as missing data. We removed individuals who had

more than 10% unanswered questions, as well as items where there were more than 10% missing responses as a quality control measure.

The remainder of the responses were subjected to Rasch analysis using the RUMM2030 Plus software (RUMM Laboratory Pty Ltd, Australia).⁷ A sample size of 50 patients was estimated to give 99% confidence with stable item calibration within ± 1.0 logits.⁸ We opened recruitment to the full POEMS clinic cohort to enable maximum numbers and disability breadth of responses.

The flowchart in Figure 1 summarises our iterative approach to scale refinement based on Rasch statistics. The initial scale was analysed for conformity to the Rasch unidimensional measurement model, to identify items which displayed the following characteristics indicating a poor fit to the Rasch model:

- Threshold disordering (based on inspection of Category Probability Curves [CPC], Figure 2). Items that did not display clear ordered differentiation of the three response options did not adequately contribute to the probability model and were removed.
- Poor individual item or person fit (based on item fit residual chi-square probability and F-statistics)
- Differential item functioning (DIF, based on inspection of Item Characteristic Curves [ICC], Figure 2). The scale should represent abilities of male and female participants and of all ages in the selected population. Where items identify functioning differently between different person characteristics (eg, men and women) these were discarded.
- Local dependency (by evaluating the overall correlation matrix, first identifying items with correlation indices above 0.9, then 0.8, then stepwise until all pairs of items with correlation indices over 0.28 were identified)
 - Each pair of items was individually inspected, to see which had the poorer item characteristic curves. The item with poorer fit was removed and overall scale characteristics were re-inspected to determine if removal improved or worsened overall model fit. If removal worsened overall model fit, the item was reinstated and the other item within the pair was removed.

Further details on this approach are presented in the Results section. Items were removed using a sequential and iterative process if they were found to exhibit one or more of these characteristics (Figure 1). During this analysis, we monitored person distribution within class intervals; overall model fit statistics and independent unidimensionality *t*-tests of the scale to determine whether each item removal improved the overall model. If item removal worsened the overall model, the item was reinstated.

2.2 | Reliability testing

Internal reliability of the scale was tested using the person separation index (PSI) (ie, that the scale is able to identify at least two groups of patients), with a PSI over 0.7 considered as acceptable reliability.

Test-retest reliability was performed by asking participants to complete the final POEMS-RODS on two separate occasions spaced

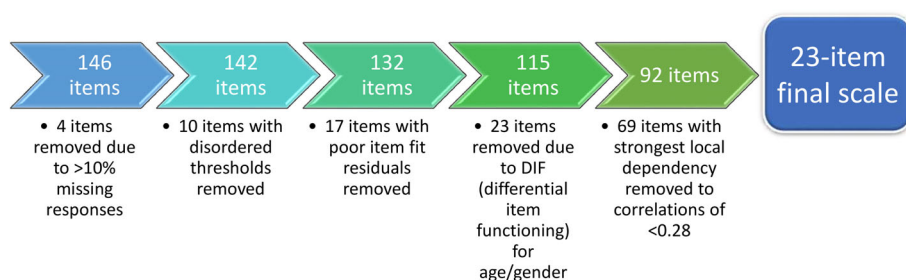


FIGURE 1 Flowchart demonstrating the sequential process of item removal to improve fit to the Rasch model

2 to 4 weeks apart. The intraclass correlation coefficient (ICC) was quantified using a two-way random effects analysis.

2.3 | Statistical analysis

RUMM2030 Plus software⁷ was used for Rasch analysis using the partial credit model. Further analysis was performed using R v4.0.2/v4.0.3 (R Core Team).⁹

2.4 | Ethics

The protocol was reviewed and approved by National Health Service (NHS) England Health Research Authority (HRA) and Research Ethics Committee (REC; reference 21/LO/0518). All participants provided written informed consent.

3 | RESULTS

3.1 | Participant characteristics and response rate

Forty-nine of 70 individuals approached from the UK national POEMS cohort, with a confirmed diagnosis of POEMS and under ongoing follow-up, completed the preliminary RODS questionnaire (response rate 70%). Median age was 59 years (SD = 12.7, range 39–81 years). 67.3% of the participants were female.

3.2 | Data quality control

One individual was removed because of a greater than 10% rate of unanswered questions. Four questions were removed from the 146-item question bank because there were missing responses from more than 10% of participants.

3.3 | Initial analysis

Initial Rasch analysis on the remaining 142 items revealed characteristics as detailed in Table 1. In summary, the initial 142-item unselected item model failed the criteria for a unidimensional scale.

3.4 | Iterative approach to data fitting

Individual item characteristics were examined to remove items contributing to model misfit, and overall scale characteristics were monitored to ensure each removal improved model fit, as displayed in Figure 1.

3.4.1 | Step 1 - Threshold ordering

Items in the scale should be able to distinguish individuals with differing ability levels reliably and consistently based on responses, with high ability individuals responding 'easy to perform', low ability individuals responding 'unable to perform' and individuals with middling ability responding 'able to perform with difficulty'. Cohort responses are summarised by the item CPC graph (Figure 2A,B). An example of an item with good threshold ordering is seen in Figure 2A. Where responses are not ordered, or overlapped, these result in difficulty identifying category thresholds (Figure 2B). Ten items were removed due to threshold disordering.

3.4.2 | Step 2 - Individual fit statistics

Individual person- and item-fit statistics can be examined to identify deviations from expected model scores. Seventeen items with poor item fit were sequentially removed. Person fit statistics were uniformly acceptable thus no further participants were removed.

3.4.3 | Step 3—differential item functioning (DIF)

DIF indicates that participants with differing characteristics (age and/or gender) respond differently to items, despite having equal overall ability levels (Figure 2C). Twenty-three items were removed because of DIF to age and/or gender.

3.4.4 | Step 4 - local dependency

Local dependency describes items that are linked in a way that responses on one item may depend on or not be independent of responses to another; this may place undue weight on one particular type of ability within the overall scale. A correlation index of 0.28 or greater is generally considered to signify local dependency. A total of 69 items were removed leaving no local dependency above a correlation index of 0.28.

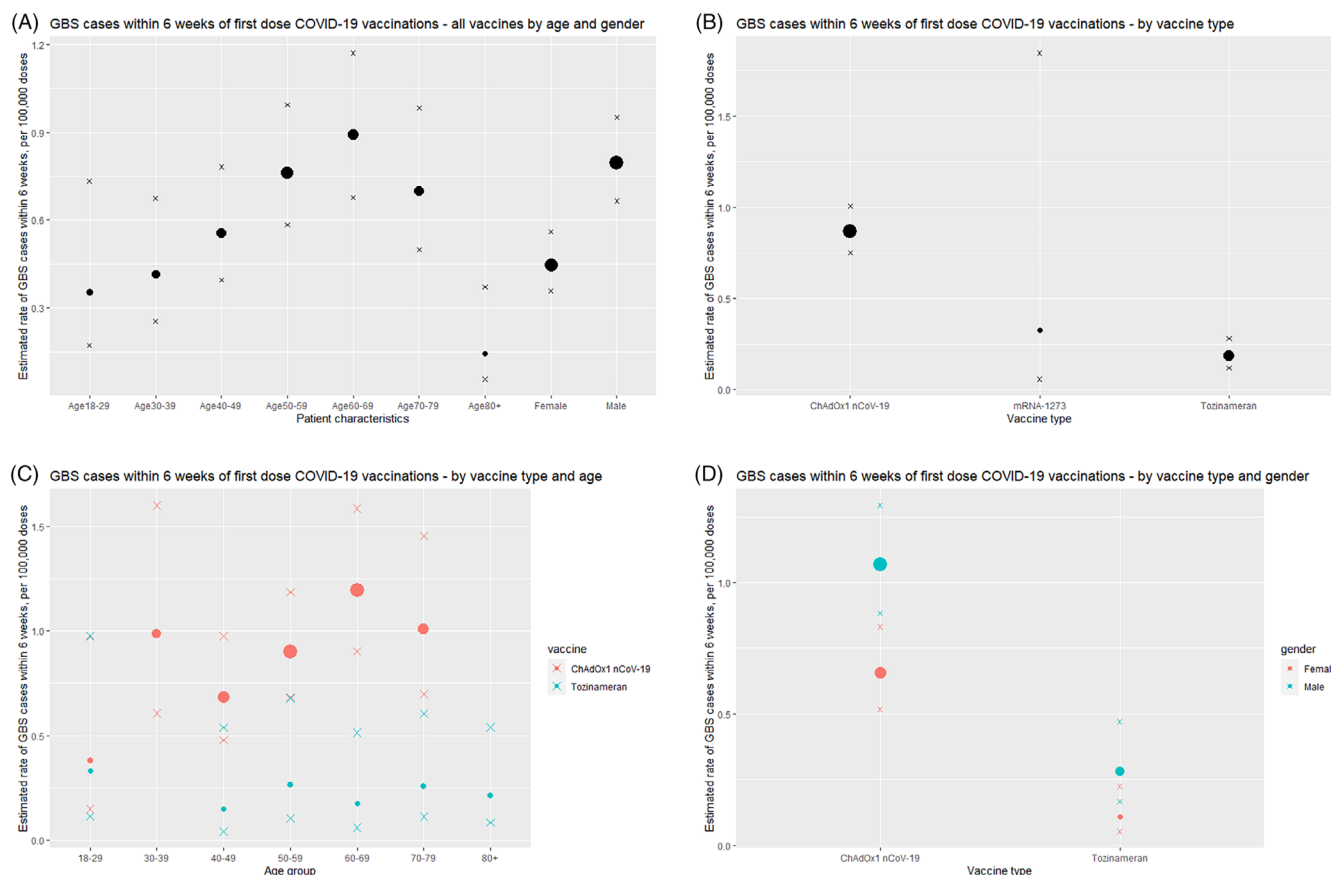


FIGURE 2 Category Probability Curves (CPC) (A, B) and Item Characteristic Curves (ICC) plotting for Differential Item Function (DIF) (C, D). In (A) and (B), the coloured lines display relative probability of participants responding 'not possible to perform' (blue), 'possible, but with some difficulty' (red), and 'possible, without any difficulty' (green) over a range of person locations (person ability on a linear scale). (A) CPC for item 'getting into a car' showing good threshold ordering - this item was retained at this point (but removed later to reduce local dependency). (B) CPC for item 'driving a car' shows disordering of thresholds, with the 'possible, but with some difficulty' category (red line) never being the most probable to be observed at any point in the scale. This item was removed. (C) ICC plot for item 'cycle a bike' shows significant DIF for gender, with male (red line) and female (blue line) participants displaying significant differences in responses at the same person location. This item was removed. (D) ICC plot for item 'put on shoes' showed no significant DIF for gender, with male and female participants responding similarly at similar levels of ability. This item was retained.

TABLE 1 Characteristics of 142-item preliminary RODS on initial Rasch analysis

Parameter	Ideal values	Observed values	Interpretation
Item fit residual	Mean = 0, SD = 1	Mean = -0.16, SD = 0.70	Acceptable fit residual
Person fit residual	Mean = 0, SD = 1	Mean = -0.27, SD = 1.41	Acceptable fit residual
Item-trait χ^2 probability	>0.05	0.12	Criteria for item invariance met
Person separation index	>0.70	0.99	Model able to differentiate >9 distinct groups
Cronbach's alpha	>0.70	0.99	Good scale reliability
Unidimensionality t -tests ^a	LL 95%CI <0.05	Proportion of significant t -tests = 0.167 (95% CI 0.105-0.228)	Scale fails criteria for unidimensionality

Abbreviations: 95% CI, 95% confidence interval; LL, lower limit.

^aUnidimensionality t -tests were performed by comparing the six most positively loaded and six most negatively loaded items on the first principal component analysis.

3.5 | Final 23-item POEMS-RODS

Following this iterative process, we produced a 23-item scale with acceptable fit and meeting Rasch criteria for unidimensionality (Table 2, Appendix S1). An online version of this is available at: <https://bit.ly/3BXZIZH>.

The threshold map and item difficulty locations for the 23 items of the final POEMS-RODS are depicted in Figure 3. Item difficulty ranged from -3.06 to +4.30 logits, with 'take a lift/elevator' being the easiest item and 'jump' being the most difficult item on the scale. The spread of item difficulty covers the range of participant ability

TABLE 2 Characteristics of 23-item preliminary RODS on Rasch analysis

Parameter	Ideal values	Observed values	Interpretation
Item fit residual	Mean = 0, SD = 1	Mean = 0.055, SD = 0.85	Good model fit
Person fit residual	Mean = 0, SD = 1	Mean = -0.186, SD = 0.874	Good model fit
Item-trait χ^2 probability	>0.05	0.982	Criteria for item invariance met
Person separation index	>0.70	0.958	Model able to differentiate >5 distinct groups
Cronbach's alpha	>0.70	0.964	Good scale reliability
Unidimensionality <i>t</i> -tests ^a	LL 95%CI <0.05	Proportion of significant <i>t</i> -tests = 0.104 (95% CI 0.043-0.166)	Scale passes the test of unidimensionality

Abbreviations: 95% CI, 95% confidence interval; LL, lower limit.

^aUnidimensionality *t*-tests were performed by comparing the six most positively loaded and six most negatively loaded items on the first principal component analysis.

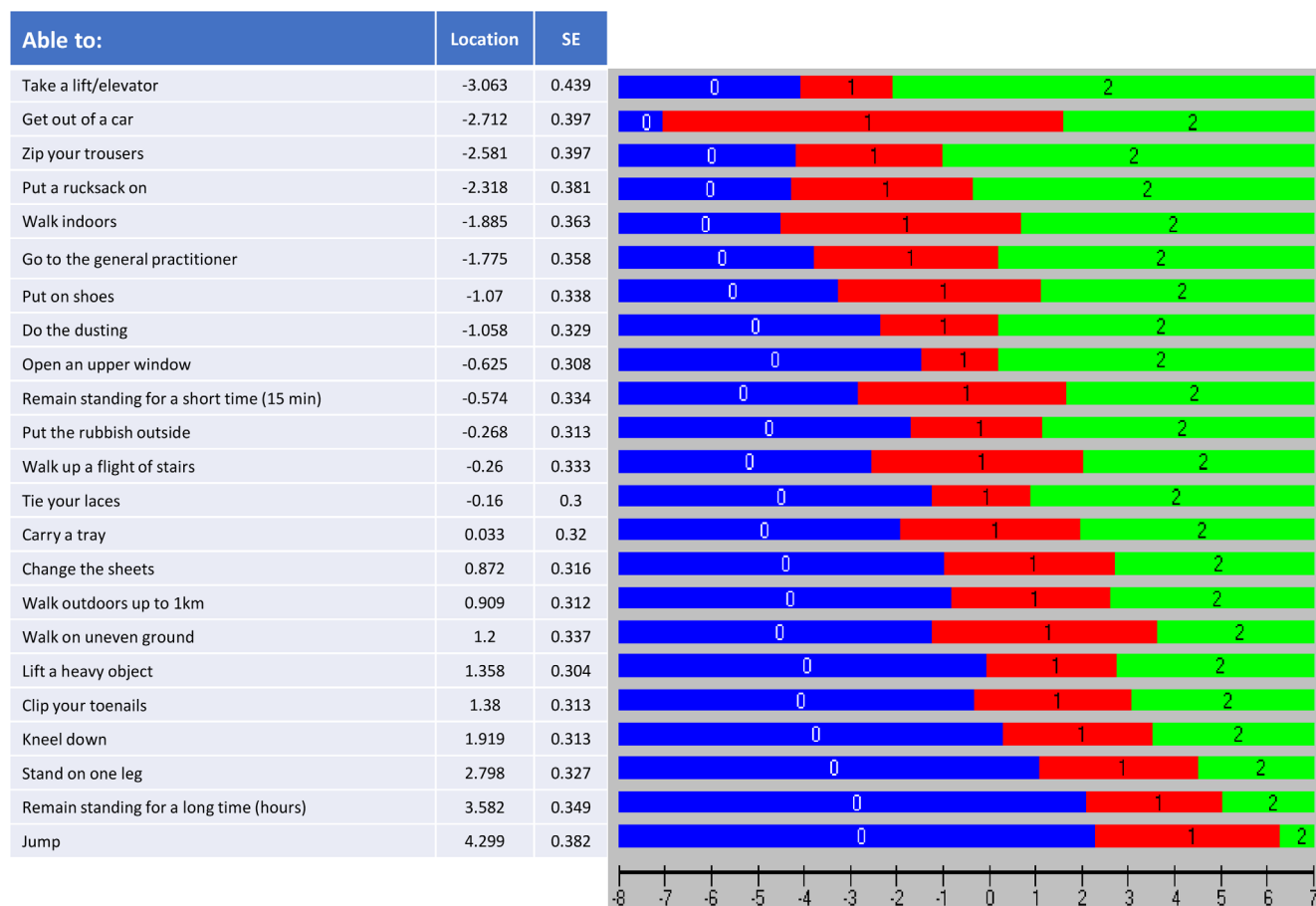


FIGURE 3 Threshold map and item difficulty locations for the 23-item final POEMS-RODS. Coloured bars correspond to participant responses of 'not possible to perform' (blue), 'possible, but with some difficulty' (red), and 'possible, without any difficulty' (green) for each item. SE: Standard error of item location (logits)

and ranges from -6.4 to +6.2 logits (Figure 4, Table 3). The raw final POEMS-RODS range from 0 to 46. For ease of interpretation, we converted the scale into a linear centile score with values ranging from 0 (most severe disability) to 100 (no disability) (Table 4).

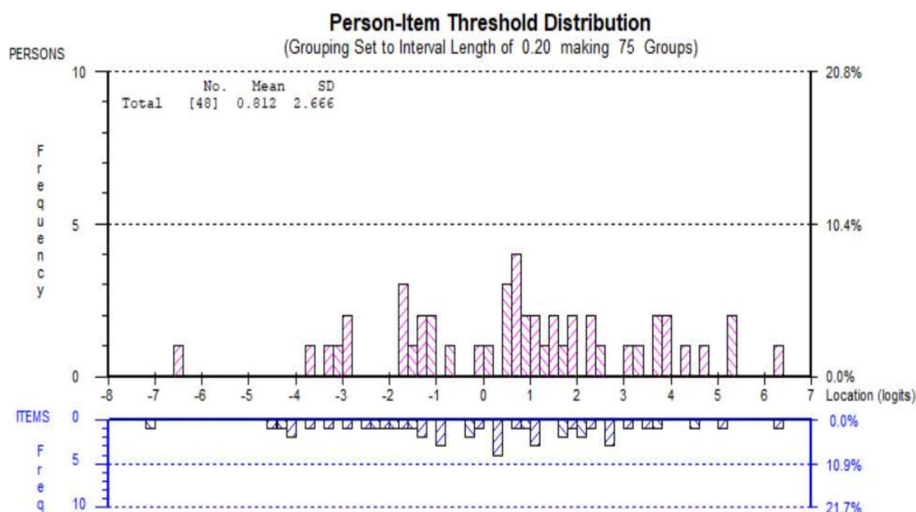
The SE for the POEMS-RODS varied with differing levels of participant ability with a 'U'-shaped pattern (Figure 5), with SE being lowest in the middle of the scale (moderate disability), similar to other RODS.^{5,10}

None of the participants scored the minimum or maximum score for the final POEMS-RODS, suggesting negligible floor and ceiling effects.

3.6 | Reliability studies

On repeat testing of the final POEMS-RODS, a PSI of 0.959 indicated that internal reliability of the scale remained good. We also demonstrated good

FIGURE 4 Person-item threshold map of the 23-item final POEMS-RODS reveals good distribution of item difficulty (lower half) relative to person ability (upper half)



test-retest reliability, with item hierarchy within the 23-item scale demonstrating an ICC of 0.936 (95% CI 0.856-0.972) and patient location an ICC of 0.954 (95% CI 0.914-0.976) (Figure 6).

4 | DISCUSSION

We developed a 23-item POEMS-RODS disability score using the UK national POEMS cohort, which conforms to Rasch statistical parameters and demonstrates good test-retest reliability.

The disability of POEMS is primarily neuropathic.⁶ However, POEMS is a multi-system disease, and in practice, other factors such as pedal oedema and cardiorespiratory aspects of the disease may also contribute towards patient disability in POEMS. Where neuropathy is minimal, which is frequently the case in Castleman disease-associated POEMS, systemic factors can account for the entirety of patient disability, although this disability is often less than cases with additional neuropathy. Overall, disability scales such as this POEMS-RODS are likely to reflect the combined multi-system disability seen in most patients with POEMS, especially at the more physically demanding end of the scale with items such as 'Jump' and 'Remain standing for a long time (hours)', which require a combination of good neurological and cardiovascular function.

Our 23-item POEMS-RODS shows good internal validity, evidenced by a high PSI of 0.959. In the construction of disability scales, external validity may be supported by correlating it with other existing scales validated for the condition; for example, the i-RODS was shown to correlate well with the Overall Disability Sum Score (ODSS).^{4,11} However, the ability to validate the scale externally is reliant on a pre-existing scale for the condition; neither the ODSS nor the modified Overall Neuropathy Limitations Scale,¹² or any other neuropathy disability scales have been validated in POEMS. Future longitudinal and responsiveness assessments will explore comparative performance to some of these non-validated scores.

The POEMS-RODS SE differs with different levels of patient ability in a characteristic 'U'-shaped curve (Figure 5), with the SE being

highest at extremes of patient ability and lowest where patient ability is middling. This is similar to previous RODS^{5,10} and reflects a more limited number of individual responses available at extremes of patient ability during scale construction. The responses of the UK POEMS cohort used to construct this scale revealed very few individuals with particularly low ability levels, as seen by the gaps on the left of the curve in Figure 5.

The uncertainty inherent in scores at the extremes of patient ability is particularly important for the interpretation of the centile scale presented in Table 4 and the interpretation of a minimal clinically important difference (MCID). Although changes in a single absolute point in the middle of the scale reflect only a 1 to 2 centile change in the linear centile score of the POEMS-RODS compared to 8-10 centiles at extreme ends, the smaller SE in the middle of the scale indicates that a small change in score in the middle of the scale is likely to be more certain, and to an extent thus more clinically important, than a larger change at the high and low ends of the scale. Because of this, although we have presented conversion to a centile form of the scale as has been done for other scales, we advise caution in interpreting changes at the extreme ends of the scale as being highly significant. Further studies of the POEMS-RODS are likely to be valuable in establishing a meaningful MCID at various points in the scale. This has been an ongoing task and a persistent debate for the i-RODS over its first decade of use.

Although the POEMS-RODS displayed good intraclass correlation coefficients (ICC) for person and item location on repeat testing, suggesting good test-retest reliability, a number of patients and items demonstrated changes in scale location outside the 95% CI (Figure 6, shaded areas). We recognise that longitudinal and larger collaborative international studies could improve the reliability and certainty of the POEMS-RODS, or result in small modifications that would reduce the uncertainties. However, larger patient cohorts are not currently available and this first version of POEMS-RODS has reasonable validity at this stage for use.

One of the main limitations of our study is the sample size. Previous studies creating RODS for inflammatory neuropathies^{4,5} have

TABLE 3 23-item POEMS-RODS item location (right side of the scale) relative to person ability (left), depicted on a single ruler of ability/difficulty (logits)

Persons	Location (logits)	Item location
x	7.0	
xx	6.0	
x x	5.0	Jump
xx xx x	4.0	Remain standing for a long time (hours)
x x xx	3.0	Stand on one leg
xx x xx x	2.0	Kneel down Clip your toenails; Lift a heavy object
xx xx xxxx xxx	1.0	Walk on uneven ground Walk outdoors up to 1 km; Change the sheets
x x x	0.0	Carry a tray Tie your laces Walk up a flight of stairs; Put the rubbish outside Remain standing for a short time (15 min) Open an upper window
xx xx x xxx	-1.0	Do the dusting; Put on shoes Go to the general practitioner
	-2.0	Walk indoors Put a rucksack on Zip your trousers Get out of a car
xx x x x	-3.0	Take a lift/elevator

TABLE 3 (Continued)

Persons	Location (logits)	Item location
	-4.0	
	-5.0	
x	-6.0	

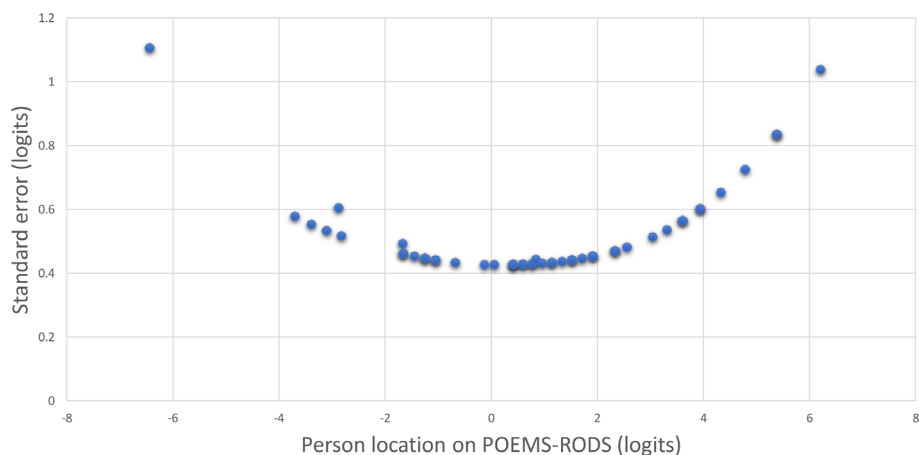
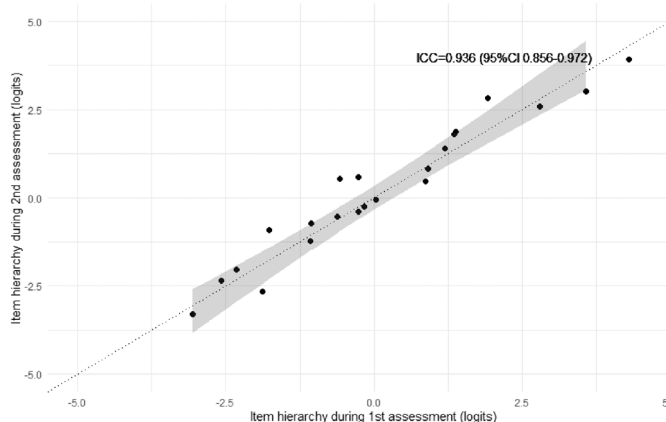
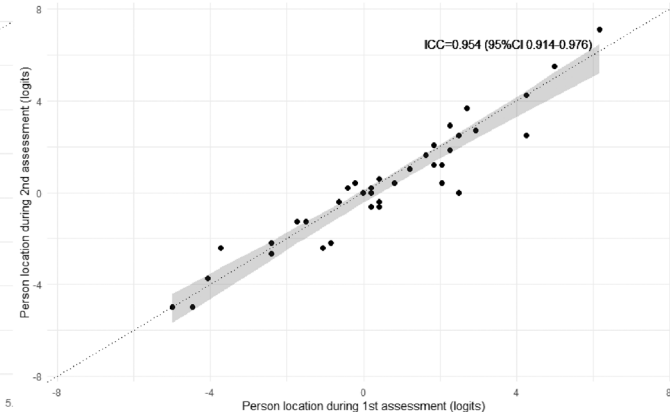
Note: x = 1 person.

TABLE 4 Nomogram conversion of 23-item POEMS-RODS from a 46-point scale to a 100-point centile score

Raw POEMS-RODS score	Rasch person location (logits)	POEMS-RODS centile score
0	-7.906	0
1	-6.433	10
2	-5.493	16
3	-4.892	20
4	-4.432	23
5	-4.044	25
6	-3.700	28
7	-3.385	30
8	-3.093	32
9	-2.820	33
10	-2.563	35
11	-2.318	37
12	-2.086	38
13	-1.863	40
14	-1.650	41
15	-1.443	42
16	-1.243	44
17	-1.047	45
18	-0.856	46
19	-0.669	47
20	-0.484	49
21	-0.301	50
22	-0.120	51
23	0.060	52
24	0.240	53
25	0.420	55
26	0.601	56
27	0.783	57

TABLE 4 (Continued)

Raw POEMS-RODS score	Rasch person location (logits)	POEMS-RODS centile score
28	0.966	58
29	1.151	59
30	1.339	61
31	1.530	62
32	1.725	63
33	1.925	64
34	2.130	66
35	2.343	67
36	2.565	69
37	2.800	70
38	3.049	72
39	3.319	73
40	3.615	75
41	3.948	78
42	4.334	80
43	4.798	83
44	5.388	87
45	6.214	92
46	7.369	100

FIGURE 5 Participant location on POEMS-RODS with corresponding SE. A U-shaped curve is seen indicating SE differs based on patient ability, with SE being lowest for participants with moderate disability.**(A)** POEMS-RODS item difficulty hierarchy (logits) between 1st and 2nd assessment**(B)** POEMS-RODS person location (logits) between 1st and 2nd assessment**FIGURE 6** Test-retest reliability of the 23-item POEMS-RODS. Shaded areas represent 95% confidence intervals. (A) Item difficulty hierarchy of the final POEMS-RODS in the first vs the second assessment. (B) Person location during first vs second assessment. ICC, intraclass correlation coefficient

utilised over 100 participants for scale design, with recommendations that more than 250 may be required for high certainty.⁸ Based on these sources, our sample size is estimated to convey a 99% confidence with stable item calibration within ± 1.0 logits.⁸ To increase the certainty of the estimations, additional validation studies with a larger cohort will be needed for the improvement of the POEMS-RODS. However, POEMS is much rarer than CIDP and MMN, and our study of 49 POEMS patients represents a large proportion of the current known UK POEMS cohort.² Because of this rarity, despite the more limited sample size, we feel that our final POEMS-RODS is reasonably representative of the range of disabilities experienced by the UK POEMS cohort.

The enrolled patient cohort for this study was heterogeneous, including patients with POEMS treated for a wide range of disease durations. Not only did this maximise sample size, but ensured the cohort reflects the range of disabilities seen in POEMS. As the study was cross-sectional, there was very little inter-examination variability. We did not explore responsiveness as most patients are highly stable, having been successfully treated, and smaller numbers are in active treatment. Future studies with a larger cohort will allow for analysis of factors (beyond age and gender) that might contribute towards Differential Item Function and heterogeneity. The disease is too rare to generate the participant numbers required for these validations.

International collaboration will be important in achieving the larger patient numbers required to confirm the broader statistical viability of the POEMS-RODS and to determine its responsiveness.¹³ Larger studies will also facilitate accurate estimations of the MCID of the POEMS-RODS, either through statistical techniques, anchoring or consensus. International engagement will also be helpful in confirming cross-cultural validity, as cultural differences are known to cause individuals of different nationalities to respond differently to certain activity items.

5 | CONCLUSION

We have built a 23-item POEMS-RODS from representative patient responses which reflects the range of disabilities within the UK POEMS cohort, conforms well to a Rasch model, and demonstrates good test-retest reliability. Larger international studies are needed to confirm the broader applicability of this scale in other countries.

ACKNOWLEDGEMENT

The authors are grateful to Dr Ingemar Merkies for his advice on Rasch scale development.

CONFLICT OF INTEREST

The authors have no conflict of interest relevant to this manuscript.

DATA AVAILABILITY STATEMENT

Anonymised data not published within this article will be made available on reasonable request by contacting the Principal Author at michaellunn@nhs.net.

ORCID

Ryan Yann Shern Keh  <https://orcid.org/0000-0001-6434-9496>

Michael P. Lunn  <https://orcid.org/0000-0003-3174-6027>

REFERENCES

- Keddie S, Foldes D, Caimari F, et al. Clinical characteristics, risk factors, and outcomes of POEMS syndrome: a longitudinal cohort study. *Neurology*. 2020;95(3):e268-e279. doi:[10.1212/WNL.0000000000009940](https://doi.org/10.1212/WNL.0000000000009940)
- Pihan M, Keddie S, D'Sa S, et al. Raised VEGF: high sensitivity and specificity in the diagnosis of POEMS syndrome. *Neurol Neuroimmunol Neuroinflamm*. 2018;5(5):e486. doi:[10.1212/NXI.0000000000000486](https://doi.org/10.1212/NXI.0000000000000486)

- Rasch G. *Probabilistic Models for some Intelligence and Attainment Tests*. Copenhagen: Danish Institute for Educational Research; 1960.
- van Nes SI, Vanhoutte EK, van Doorn PA, et al. Rasch-built overall disability scale (R-ODS) for immune-mediated peripheral neuropathies. *Neurology*. 2011;76(4):337-345. doi:[10.1212/WNL.0b013e318208824b](https://doi.org/10.1212/WNL.0b013e318208824b)
- Vanhoutte EK, Faber EG, van Nes SI, et al. Rasch-built overall disability scale for multifocal motor neuropathy (MMNRODS[©]). *J Peripher Nerv Syst*. 2015;20(3):296-305.
- Keddie S, Lunn MP. POEMS syndrome. *Curr Opin Neurol*. 2018;31(5):551-558. doi:[10.1097/WCO.0000000000000610](https://doi.org/10.1097/WCO.0000000000000610)
- Rummlab Pty Ltd. RUMM2030+. <https://www.rummlab.com.au/>. Accessed online on: 29 July 29, 2022.
- Rasch Measurement Transactions. Sample size and item calibration [or person measure] stability. <https://www.rasch.org/rmt/rmt74m.htm>. Accessed online on July 29, 2022.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing; 2020. <https://www.R-project.org/>. Accessed online on: July 29, 2022
- Vanhoutte EK, Draak THP, Gorson KC, et al. Impairment measures versus inflammatory RODS in GBS and CIDP: a responsiveness comparison. *J Peripher Nerv Syst*. 2015;20(3):289-295. doi:[10.1111/jns.12118](https://doi.org/10.1111/jns.12118)
- Merkies IS, Schmitz PIM, van der Meché FGA, Samijn JP, van Doorn P, Inflammatory Neuropathy Cause and Treatment (INCAT) group. Clinimetric evaluation of a new overall disability scale in immune mediated polyneuropathies. *J Neurol Neurosurg Psychiatry*. 2002;72(5):596-601.
- Graham RC, Hughes RA. A modified peripheral neuropathy scale: the overall neuropathy limitations scale. *J Neurol Neurosurg Psychiatry*. 2006;77(8):973-976. doi:[10.1136/jnnp.2005.081547](https://doi.org/10.1136/jnnp.2005.081547)
- Draak THP, Vanhoutte EK, van Nes SI, et al. PeriNomS Study Group. Changing outcome in inflammatory neuropathies: Rasch-comparative responsiveness. *Neurology*. 2014;83(23):2124-2132. doi:[10.1212/WNL.0000000000001044](https://doi.org/10.1212/WNL.0000000000001044)

SUPPORTING INFORMATION

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

How to cite this article: Keh RYS, Lilleker JB, Lavin T, et al. Rasch-built overall disability scale for POEMS syndrome (POEMS-RODS). *J Peripher Nerv Syst*. 2022;27(4):302-310. doi:[10.1111/jns.12516](https://doi.org/10.1111/jns.12516)