Comparative value and determinants of suitability of outcome measures in treated chronic inflammatory demyelinating polyneuropathy

Yusuf A. Rajabally,[1,2] MD, FRCP   Majid Ghasemi[1] MD.

1. **Regional Neuromuscular Service, University Hospitals Birmingham, Birmingham, U.K.**

2. **School of Life and Health Sciences & Aston Medical School, Aston University, Birmingham, U.K.**

REVISED VERSION R2

CLINICAL RESEARCH ARTICLE

Disclosures:

*"**We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines**."*

Correspondence to:
Yusuf A. Rajabally

School of Life and Health Sciences,
Aston Brain Centre,
Aston University,
Aston Triangle,
Birmingham B4 7ET, UK.
E-mail: y.rajabally@aston.ac.uk

Abstract.

Introduction: The comparative value and suitability of outcome measures are uncertain in chronic inflammatory demyelinating polyneuropathy (CIDP).

Methods: We studied 35 patients with CIDP using the Overall Neuropathy Limitation Scale (ONLS), the inflammatory Rasch-built Overall Disability Scale (I-RODS) and the Medical Research Council Sum Score (MRCSS).

Results: Significant associations were determined between initial deficit and ONLS (p=0.002) and MRCSS improvement (p=0.001), but not I-RODS. A strong inverse correlation was observed between disease duration and I-RODS (p=0.002), but not ONLS/MRCSS improvement. A strong association was observed between age ≤40 years and I-RODS (p=0.001), but not ONLS/MRCSS improvement. Using minimum important differences, ONLS and I-RODS sensitivities were comparable (p=0.19).

Discussion: ONLS and I-RODS are equally sensitive in identifying change. Ease of administration, better ability to detect improvement in severe disease, greater amplitude responses throughout the disease and in older subjects, favor the ONLS. The I-RODS appears more useful in early disease/younger patients.

Introduction.

The use of disability scales has gained importance in the literature in chronic inflammatory demyelinating polyneuropathy (CIDP) in recent years. Their use in research settings has become widespread. Although the more recently developed inflammatory Rasch-built Overall Disability Scale (I-RODS) (1) has been described as testing a wider range of activities and having greater responsiveness, it has not, to date, been used in clinical trials as a primary outcome measure. On the other hand, the similar Overall Neuropathy Limitation Scale (ONLS) (2) and Inflammatory Neuropathy Cause and Treatment (INCAT) (3) scales have remained the preferred primary outcome in research studies (4-8). The INCAT or ONLS have the advantage over the I-RODS of being easier and quicker to administer.

Disability scales are not consistently used in most non-academic institutions (9). Their comparative value remains uncertain in clinical practice. The value of score changes, and interpretation that may result from them, poses difficulty. The clinical application of the concept of minimum important difference (MID) is complex (10). The literature on the subject has been focused on methods allowing calculation in a research setting, although, to

our knowledge, neither has been directly applied in research trials as a primary outcome measure, nor in clinical practice in CIDP (10). Establishing what represents meaningful change remains challenging in the absence of an evidence-base regarding the MID. This is however highly desirable for clinical management (11).

The main aims of this analysis were to explore the determinants of change on these existing scales in clinical practice, establish MID-defined clinically-applicable cut-offs for use in therapeutic management and monitoring, and to compare scale sensitivities with these cut-offs in evaluating treatment response.

Methods.

We retrospectively reviewed electronic hospital records of all patients meeting clinical and electrophysiological criteria for "definite" or "probable" CIDP as per European Federation of Neurological Societies/Peripheral Nerve Society (EFNS/PNS) Guidelines (12), attending the Inflammatory Neuropathy Clinic at Queen Elizabeth Hospital, University Hospitals of Birmingham, U.K., between June 2014 and December 2018. Scales required for inclusion in this study were the Overall Neuropathy Limitation Scale (ONLS), inflammatory Rasch-built Overall Disability Scale (I-RODS) and the Medical Research Council (MRC) sum score (MRCSS). Patients who had been assessed by these 3 scales and who had displayed any degree of improvement on any of the 3 with therapeutic intervention, were considered as "responders" for the purposes of this study and were selected for further analysis. We decided on this definition despite the limitations in view of the uncertainties of MIDs in our cohort prior to our study and as improvement is defined variably in clinical practice.

For each patient, we identified demographics, diagnostic criteria category fulfilment, CIDP subtype and treatment given. All scales had been administered pre- and post-therapeutic intervention. In the case of de-novo patients, assessments took place between 2 and 4 weeks before and after first treatment. Similarly, in all other cases, patients were assessed (i) 2 to 4 weeks preceding and following re-initiation of treatment after prior withdrawal resulting in deterioration, or (ii) 2 to 4 weeks preceding and following re-increase of treatment dose after a prior dose reduction which had been followed by deterioration. The MRCSS (out of 80) was determined by adding individual scores (0 to 5 for each muscle group) from 8 muscle groups bilaterally, including shoulder abductors, elbow flexors, wrist extensors, finger abductors, hip flexors, knee flexors, ankle dorsiflexors and extensor hallucis. The ONLS score ranging from 0 (best) to 12 (worst) was ascertained by physician-administered interview, combining upper limb (out of 5) and lower limb (out of 7) scores. Upper limb ONLS grades range from 0 (no symptoms or disability), to 5 (inability to perform purposeful arm movements) and lower limb ONLS grades range from 0 (ability to run or climb stairs without difficulty), to 7 (inability to perform purposeful leg movements) (2). The I-RODS raw score ranging from 0 (worst) to 48 (best) was determined by physician-administered interview of the 24 items on that scale, rating inability (0/2), difficulty with (1/2), or normality (2/2) for each task (1).

Mean improvements on each scale and standard deviations (S.D.) were determined. Intercorrelation studies of the MRCSS with ONLS and I-RODS were performed, for initial and final measures as well as improvement levels. We attempted to ascertain eventual determinants of the amplitude of response for each scale, amongst several co-variates

including disability at onset as measured on the studied scale, disease duration, age at onset, age at time of analysis and gender. In addition, we calculated, for our cohort, the MID, using a validated distribution-based statistical method with ½ S.D., which has been described as showing equivalence to anchor-based methods (13, 14). Furthermore, as a distribution-based method, ½ S.D. has been found to be a widely valid estimate of the MID across multiple studies (15). Responder rates achieved in relation to the closest measurable scores above the calculated MID for each scale, were established. We subsequently attempted to determine the most sensitive disability scale for use in the clinical setting, with the established MID. This was performed by re-defining responder status for each scale using the MID as cut-off and calculating the sensitivity in each case (number reaching responder status as per new definition/35 x 100).

Comparison of proportions were performed with Fisher Exact tests and comparisons of means by independent t-tests. Correlations were performed using Pearson's correlation. Bonferroni correction was applied for multiple comparisons. Correlations levels were defined as "perfect", "very strong", "moderate", "fair", "poor", or "none", as previously suggested for medical research (16). Multiple regression was used to ascertain eventual independent associations. Statistical analyses were performed using SPSS 25.0 software.

This analysis was performed as part of a retrospective Clinical Audit approved and registered at University Hospitals of Birmingham, U.K. on evaluation of the use of outcome measures for inflammatory neuropathy in our Inflammatory Clinic (CARMS 14615, September 2018).

Results.

We recruited 35 patients in the current study. The main characteristics of the studied cohort are summarised in Table 1.

The mean initial MRCSS was 62.29 (S.D.: 8.32), mean initial ONLS was 4.54 (S.D.: 1.42) and mean initial I-RODS was 26.26 (S.D.: 8.47). Mean final MRCSS was 73.57 (S.D.: 7.27),

mean final ONLS was 2.51 (S.D.: 1.48) and mean final I-RODS was 34.94 (S.D.: 8.92). Mean MRCSS improvement was 11.20 (S.D.: 6.35; range: 2-29), mean ONLS improvement was 2.03 (S.D.: 1.54; range: 0-7) and mean I-RODS improvement was 8.66 (S.D.: 6.87; range: 0-29).

Initial MRCSS correlated fairly with initial ONLS scores (r=-0.526; p=0.001) and with initial I-RODS scores (r=0.553; p=0.001). Initial ONLS and I-RODS scores were also fairly inter-correlated (r=-0.484; p=0.003). Final MRCSS correlated moderately with final ONLS scores (r=-0.662; p<0.001) and I-RODS (r=0.678; p<0.001). Final ONLS and I-RODS scores were very strongly correlated (r=-0.822; p<0.001).

The improvement in MRCSS and ONLS scores were moderately correlated (r= 0.633; p<0.001), as were the improvements in ONLS and I-RODS scores (r=0.664; p<0.001). A fair correlation was ascertained between the MRCSS improvement and the I-RODS improvement (r=0.412; p=0.014).

Table 2. summarises results of correlation studies with the 5 studied potential determinants of responses for each of the 3 scales (Bonferroni-corrected; significance p<0.01). Fair correlations were ascertained between pre-treatment deficit and the improvement achieved for the MRCSS and the ONLS, but none was found for the I-RODS. A fair inverse correlation was ascertained between disease duration and I-RODS improvement but none was found for the other 2 scales. Age of onset did not correlate with improvement on any of the 3 scales. A fair correlation was found between age ≤40 years and I-RODS improvement but

none observed for the other 2 scales. Gender did not correlate with improvement on any scale.

Multiple linear regression demonstrated that I-RODS improvement was independently inversely associated with disease duration (p=0.005) and age <40 years at time of analysis (p=0.002).

The MID determined with the distribution method using the ½ S.D., was 3.18 for MRCSS, 0.77 for ONLS and 3.44 for I-RODS. Using corresponding minimal clinically-applicable cut-offs, i.e. 4 points for the MRCSS, 1 point for the ONLS and 4 points for the I-RODS, the sensitivity of the MRCSS was 94.3% (33/35), that of the ONLS was of 91.4% (32/35), and that of the I-RODS, of 77.1% (27/35), in establishing treatment response. The differences were not significant (MRCSS vs. ONLS: p=1; MRCSS vs. I-RODS: p=0.084; ONLS vs. I-RODS: p=0.19).

Discussion.

We studied 3 currently commonly used clinical scales in CIDP management. The MRCSS is exclusively muscle strength-related and has been used in clinical trials and practice for several decades. Recent studies have challenged its validity and questioned its reliability (17). It remains however in its current form of six-point testing, as utilised in our unit, still commonly used in practice. The ONLS was derived from the Guy's disability scale which also in the interim led to use of the INCAT (Inflammatory Neuropathy Cause and Treatment) scale (2). Despite poor content and structural validity, the INCAT/ONLS are considered reliable and responsive (18). The minimum detectable change with the INCAT was however found greater than the derived MID from patients participating in a research study (19), suggesting suboptimal sensitivity to detect improvement in some subjects, compounded by inadequately low weight attributed to proximal arm function and stamina-related activities (18). In more recent years, the inflammatory Rasch-built Overall Disability Scale (I-RODS) has been developed (1). Compared to the ONLS, the I-RODS offers assessment of a wider range of item difficulties, and therefore evaluation of patients of different ability. It has, as a result, been adopted in daily practice in many centres.

We found high levels of correlation between MRCSS and ONLS improvement levels and between ONLS and I-RODS changes. Only moderate correlation was found between MRCSS and I-RODS and this is consistent with the latter scale exploring non-strength-related disabilities particularly stamina-related tasks and sensory dysfunction.

We ascertained that greater impairment at baseline was more likely to be followed by larger score improvements for the MRCSS and the ONLS but not I-RODS. This represents a benefit of the former 2 scales, as allows easier recognition of amelioration. We found a significant inverse association of disease duration with I-RODS improvement level. No association was ascertained with ONLS or MRCSS improvement. The I-RODS therefore appeared to demonstrate greater score changes and be of greater value, in early disease. These findings suggest the greater usefulness of the I-RODS in early CIDP with shorter disease duration but also conversely, that the ONLS and MRCSS are equally useful in all disease stages, including later in its course. It is possible that the I-RODS shows more change in early disease as larger differences may be observed on treatment initiation, rather than later when treatment changes are done, when careful attention is paid to change for decisions to re-introduce treatment or increase doses.

In relation to age, we found a strong association between age ≤40 years at time of evaluation and I-RODS improvement. There was however no correlation with the degree of ONLS or MRCSS improvement. This suggests limited practical clinical value of the I-RODS in older patients with CIDP, the mean age of our cohort of 58.74 years being similar to those of previously described populations, and in keeping with CIDP being a disease more common in older subjects (20-22). Intuitively, the use of the I-RODS, in view of the range of disability items considered, would be likely more helpful in younger and therefore often more active subjects, and our results support this impression. The independent associations we found between disease duration and age <40 years at time of study and the I-RODS score improvement, highlight the impact of both of these factors individually on the value of this

scale. Variability over time has been described with Patient Reported Outcome Measures (PROM) in other disorders such as ankylosing spondylitis (23). In addition, trends for greater improvement in younger subjects has been reported with other PROMs, such as the Oxford Knee Score, after total knee replacements (24). It is likely that age and the timing of assessment are of similar importance in CIDP, and that the current "one scale for all and at all times" approach to CIDP clinical monitoring and management, may be unsuitable.

The MIDs determined from our cohort were comparable to those previously published from research studies for both the ONLS (comparing to published INCAT score MID) (19) and I-RODS (10). We found that using these MIDs, the ONLS had comparable sensitivity to the I-RODS in identifying clinically-meaningful improvement. The MID for MRCSS found in our subjects (3.18) was otherwise comparable to that found both with the distribution method using ½ S.D. (3.53) and with the anchor-based approach, in a research setting (3.07) (10).

Our study has several limitations. The number of patients studied was relatively small and from a single centre. The population studied comprised patients of different CIDP subtypes including LSS and newly described variants with anti-paranodal antibodies (25) and was therefore not homogeneous. This methodology is however consistent with that of previous studies (10) and is consistent with CIDP cohort distribution as encountered in clinical practice. We opted to include the different CIDP subtypes as atypical forms may evolve progressively with time into the typical phenotype as recently demonstrated (26), which may make rigorous separation of the subtypes illusory. Also, the administration of scales was performed by a clinician/interviewer familiar with the scales, which may not be the routine

practice of many units, where the I-RODS is often completed by patients themselves. Finally, the use of distribution methods with ½ S.D. for ascertainment of MID, although previously found equivalent to the anchor-based technique, was not validated by the latter in our study.

In conclusion, the ONLS and the I-RODS, using clinical practice-derived MIDs, appear of comparable sensitivity in identifying change in treated CIDP in clinical practice. This challenges the impression of greater adequacy of more detailed evaluations in all subjects and at all times. It is noteworthy that our findings appear consistent with the detailed results of a recent phase 3 trial of subcutaneous immunoglobulin for CIDP (8). In this trial, although 58.8% of the placebo group relapsed as defined by a modified-INCAT score deterioration of 1 point, with a median change from baseline of 1.0 point, the median change of the percentile I-RODS score was only of -3.0 (1), and therefore smaller than a MID of 4 raw points, as found previously (10), and confirmed by our current study. Ease of administration with, an albeit non-significant, better relative sensitivity, better ability to detect improvement in more severely affected subjects, greater value throughout the disease and older subjects, appear favour the ONLS in CIDP, which is a long-term chronic illness, more prevalent with age. The I-RODS appears more helpful early in the disease and in younger patients. The MRCSS, despite its described and recognised limitations, may in view of its benefits in late disease stages and older patients, remain of potential benefit in the global evaluation, although it cannot clearly, in isolation, be the sole driver of therapeutic decisions. We believe the MRCSS may also help to differentiate CIDP-related deficits to those resulting from other non-neurological causes producing restrictions of mobility and function, although this requires further study, as conversely, pain from any cause may alter MRC ratings. This is,

again, especially relevant in a disease more frequent with age, and therefore, often associated with co-morbidities.

Further larger multicentre prospective studies of clinical practice are needed to determine the precise place of scale choice in the assessment of different patients, with different types of CIDP, at different times in their disease course.

Table 1.: Demographic Characteristics of 35 patients with treated CIDP

Characteristic	Mean ± SD or Number (%)
Mean age	58.74 (±13.74) years
Mean age at onset	52.63 (±14.10) years
Mean disease duration	71.94 (±72.35) months
Gender	
Male	24 (68.5%)
Female	11 (31.5%)

EFNS/PNS Guidelines Diagnostic category	
Definite	31 (88.6%)
Probable	4 (11.4%)
CIDP Subtype	
Typical	30 (85.7%)
Lewis-Sumner Syndrome	4 (11.4%)
Pure motor	1 (2.9%)
Treatment	
IVIg	28 (80%)
IV Methylprednisolone	5 (14.3%)
Rituximab	2 (5.7%)

Abbreviations:

CIDP: chronic inflammatory demyelinating polyneuropathy

EFNS/PNS: European Federation of Neurological Societies/Peripheral Nerve Society

Table 2. Correlations between improvements in MRCSS, ONLS and I-RODS and pre-treatment deficit, disease duration and age <40 years.

	Pre-treatment deficit	Disease Duration	Age <40 years at time of study	Age at disease onset	Gender

MRCSS improvement	r=-0.525; *p=0.001*	NS	NS	NS	NS
ONLS improvement	r=0.502; *p=0.002*	NS	NS	NS	NS
I-RODS improvement	NS	r= -0.502; *p=0.002*	r=0.551; *p=0.001*	NS	NS

r = Pearson's correlation; NS: non-significant- Bonferroni correction applied- significance: p<0.01

Abbreviations:

MRCSS: Medical Research Council Sum Score

ONLS: Overall Neuropathy Limitation Scale

I-RODS: Inflammatory Rasch-Built Overall Disability Scale

References

1.   van Nes SI, Vanhoutte EK, van Doorn PA, Hermans M, Bakkers M, Kuitwaard K, et al. Rasch-built Overall Disability Scale (R-ODS) for immune-mediated peripheral neuropathies. Neurology. 2011;76(4):337-45.

2.   Graham RC, Hughes RA. A modified peripheral neuropathy scale: the Overall Neuropathy Limitations Scale. Journal of neurology, neurosurgery, and psychiatry. 2006;77(8):973-6.

3.   Hughes R, Bensa S, Willison H, Van den Bergh P, Comi G, Illa I, et al. Randomized controlled trial of intravenous immunoglobulin versus oral prednisolone in chronic inflammatory demyelinating polyradiculoneuropathy. Annals of neurology. 2001;50(2):195-201.

4.   Randomised controlled trial of methotrexate for chronic inflammatory demyelinating polyradiculoneuropathy (RMC trial): a pilot, multicentre study. The Lancet Neurology. 2009;8(2):158-64.

5.   Hughes RA, Donofrio P, Bril V, Dalakas MC, Deng C, Hanna K, et al. Intravenous immune globulin (10% caprylate-chromatography purified) for the treatment of chronic inflammatory demyelinating polyradiculoneuropathy (ICE study): a randomised placebo-controlled trial. The Lancet Neurology. 2008;7(2):136-44.

6.   Kuwabara S, Mori M, Misawa S, Suzuki M, Nishiyama K, Mutoh T, et al. Intravenous immunoglobulin for maintenance treatment of chronic inflammatory demyelinating polyneuropathy: a multicentre, open-label, 52-week phase III trial. Journal of neurology, neurosurgery, and psychiatry. 2017;88(10):832-8.

7.   Nobile-Orazio E, Cocito D, Jann S, Uncini A, Beghi E, Messina P, et al. Intravenous immunoglobulin versus intravenous methylprednisolone for chronic inflammatory demyelinating polyradiculoneuropathy: a randomised controlled trial. The Lancet Neurology. 2012;11(6):493-502.

8.   van Schaik IN, Bril V, van Geloven N, Hartung HP, Lewis RA, Sobue G, et al. Subcutaneous immunoglobulin for maintenance treatment in chronic inflammatory demyelinating polyneuropathy (PATH): a randomised, double-blind, placebo-controlled, phase 3 trial. The Lancet Neurology. 2018;17(1):35-46.

9.   Rajabally YA. Assessing the benefit of treatment in chronic inflammatory demyelinating polyneuropathy: the challenges of clinical practice. Neurodegenerative disease management. 2018;8(5):285-8.

10.   Merkies IS, van Nes SI, Hanna K, Hughes RA, Deng C. Confirming the efficacy of intravenous immunoglobulin in CIDP through minimum clinically important differences:

shifting from statistical significance to clinical relevance. Journal of neurology, neurosurgery, and psychiatry. 2010;81(11):1194-9.

11.   Rajabally YA, Afzal S. Clinical and economic comparison of an individualised immunoglobulin protocol vs. standard dosing for chronic inflammatory demyelinating polyneuropathy. Journal of neurology. 2019;266(2):461-7.

12.   Van den Bergh PY, Hadden RD, Bouche P, Cornblath DR, Hahn A, Illa I, et al. European Federation of Neurological Societies/Peripheral Nerve Society guideline on management of chronic inflammatory demyelinating polyradiculoneuropathy: report of a joint task force of the European Federation of Neurological Societies and the Peripheral Nerve Society - first revision. European journal of neurology. 2010;17(3):356-63.

13.   Eton DT, Cella D, Yost KJ, Yount SE, Peterman AH, Neuberg DS, et al. A combination of distribution- and anchor-based approaches determined minimally important differences (MIDs) for four endpoints in a breast cancer scale. Journal of clinical epidemiology. 2004;57(9):898-910.

14.   Norman GR, Sridhar FG, Guyatt GH, Walter SD. Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. Medical care. 2001;39(10):1039-47.

15.   Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Medical care. 2003;41(5):582-92.

16.   Chan YH. Biostatistics 104: correlational analysis. Singapore medical journal. 2003;44(12):614-9.

17.   Vanhoutte EK, Faber CG, van Nes SI, Jacobs BC, van Doorn PA, van Koningsveld R, et al. Modifying the Medical Research Council grading system through Rasch analyses. Brain : a journal of neurology. 2012;135(Pt 5):1639-49.

18.   Breiner A, Barnett C, Bril V. INCAT disability score: a critical analysis of its measurement properties. Muscle & nerve. 2014;50(2):164-9.

19.   Eftimov F, van Schaik I. Chronic inflammatory demyelinating polyradiculoneuropathy: update on clinical features, phenotypes and treatment options. Current opinion in neurology. 2013;26(5):496-502.

20.   Laughlin RS, Dyck PJ, Melton LJ, 3rd, Leibson C, Ransom J, Dyck PJ. Incidence and prevalence of CIDP and the association of diabetes mellitus. Neurology. 2009;73(1):39-45.

21.   Mahdi-Rogers M, Hughes RA. Epidemiology of chronic inflammatory neuropathies in southeast England. European journal of neurology. 2014;21(1):28-33.

22.   Rajabally YA, Simpson BS, Beri S, Bankart J, Gosalakkal JA. Epidemiologic variability of chronic inflammatory demyelinating polyneuropathy with different diagnostic criteria: study of a UK population. Muscle & nerve. 2009;39(4):432-8.

23.   Essers I, Boonen A, Busch M, van der Heijde D, Keszei AP, Landewe R, et al. Fluctuations in patient reported disease activity, pain and global being in patients with ankylosing spondylitis. Rheumatology (Oxford, England). 2016;55(11):2014-22.

24.   Williams DP, Price AJ, Beard DJ, Hadfield SG, Arden NK, Murray DW, et al. The effects of age on patient-reported outcome measures in total knee replacements. The bone & joint journal. 2013;95-b(1):38-44.

25.   Querol L, Devaux J, Rojas-Garcia R, Illa I. Autoantibodies in chronic inflammatory neuropathies: diagnostic and therapeutic implications. Nature reviews Neurology. 2017;13(9):533-47.

26.   Doneddu PE, Cocito D, Manganelli F, Fazio R, Briani C, Filosto M, et al. Atypical CIDP: diagnostic criteria, progression and treatment response. Data from the Italian CIDP Database. Journal of neurology, neurosurgery, and psychiatry. 2019;90(2):125-32.