Received: 24 February 2022 | Revised: 5 April 2022 | Accepted: 10 April 2022

DOI: 10.1111/jns.12497

RESEARCH REPORT

WILEY

Assessing deterioration using impairment and functional outcome measures in chronic inflammatory demyelinating polyneuropathy: A post-hoc analysis of the immunoglobulin overtreatment in CIDP trial

Robin van Veen[1] | Luuk Wieske[1,2] | Ilse Lucke[1] | Max E. Adrichem[1] | Ingemar S. J. Merkies[3,4] | Ivo N. van Schaik[1,5] | Filip Eftimov[1]

[1]Department of Neurology, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

[2]Department of Clinical Neurophysiology, St Antonius Hospital, Nieuwegein, The Netherlands

[3]Maastricht Academic Medical Centre, Maastricht, The Netherlands

[4]Curaçao Medical Centre, Willemstad, Curacao

[5]Spaarne Gasthuis, Haarlem, The Netherlands

Correspondence
Filip Eftimov, Department of Neurology, Amsterdam UMC, University of Amsterdam, Meibergdreef 9, 1100DD Amsterdam, The Netherlands.
Email: f.eftimov@amsterdamumc.nl

Funding information
Dutch Governmental grant (ZonMw), Grant/Award Number: 836021010; Prinses Beatrix Spierfonds; Sanquin Plasma Products B.V.

Abstract

It is unclear whether frequently used cutoff values for outcome measures defining minimal clinically important differences (MCIDs) can accurately identify meaningful deterioration in chronic inflammatory demyelinating polyneuropathy (CIDP). We used data from the immunoglobulin overtreatment in CIDP (IOC) trial, in which 60 clinically stable patients with CIDP were randomized to intravenous immunoglobulin (IVIg) withdrawal or continuation. We calculated change scores of the Inflammatory Rasch-Built Overall Disability Scale (I-RODS), grip strength, and Medical Research Council-sum score (MRC-SS) and classified visits based on a treatment anchor (ie, decision to restart/increase treatment after reaching a predefined early endpoint of deterioration). The variability of scores in patients without deterioration was calculated using the limits of agreement. We defined optimized MCIDs for deterioration and specific combinations of MCIDs from different outcome measures, and subsequently calculated the accuracies of the (combined) MCIDs. Substantial variability was found in scores of the I-RODS, grip strength and MRC-SS in patients without deterioration over time, and most MCIDs were within the limits of the variability observed in patients without deterioration. Some MCID cut-offs were insensitive but highly specific for detecting deterioration, for example, the MCID-SE of −1.96 of the I-RODS and −2 point on the MRC-SS. Others were sensitive, but less specific, for example, −4 centiles of the I-RODS. Some combined MCIDs resulted in high specificities and moderate sensitivities. Our results suggest that clinically important deterioration cannot be distinguished from variability over time with currently used MCIDs on the individual level. Combinations of MCIDs might improve the accuracy of determining deterioration, but this needs validation.

KEYWORDS
chronic inflammatory demyelinating polyneuropathy, minimum important difference, outcome measures

1 | INTRODUCTION

Outcome measures assessing changes in disability or impairment are the preferred way to evaluate outcome in chronic inflammatory demyelinating polyneuropathy (CIDP) in clinical trials and clinical care.[1,2] However, it is often difficult to interpret whether changes measured on these outcome measures represent clinical meaningful improvement or deterioration, especially if changes are small. This implies that, in clinical practice, a patient who shows a small change on an outcome measure will generally be judged as showing a non-meaningful change. To improve interpretability of change scores, the concept of a minimal clinically importance difference (MCID) has been introduced to define if a change on a certain outcome measure can be perceived as relevant.[3] Generally, an MCID can be based on an anchor-based method, reflecting the patient's or physician's perspective, or a distribution-based method, reflecting statistical properties.[4] Although the ideal approach to MCID is still a matter of debate, it is often recommended to base an MCID primarily on anchors, as this would better reflect clinical relevance.[4] Other authors advocate to use both an anchor-based MCID as well as a distribution-based MCID.[5] Ideally, the cutoff of an anchor-based MCID should be larger than variability in a stable population in order to distinguish clinically important change on an outcome measure from random variability.[6] In addition, MCIDs are not static and may vary between clinical scenarios like the population, disease severity and the direction of change. For example, the MCID for deterioration may be smaller than the MCID for improvement.[7] In case of CIDP, a small, negative change while tapering treatment might be considered meaningful by a patient, while the same positive change after start of induction treatment might be considered as non-meaningful.

The most important outcome measures in CIDP include the Inflammatory Neuropathy Cause and Treatment (INCAT) disability scale and the linearly weighed patient-reported Inflammatory-Rasch-built Overall Disability Scale (I-RODS),[1,2] both assessing disability, and grip strength and the Medical Research Council-sum score (MRC-SS), assessing impairment. In CIDP, MCID cutoffs for all these outcome measures have been calculated, mostly based on distribution-based methods and only for the clinical context of improvement.[8-13] Only two small studies have focussed on anchor-based MCIDs for improvement, for the INCAT disability scale, grip strength and the MRC-SS, based on a single external anchor.[8,14] Despite these limitations, the current MCID cutoffs are used in clinical (randomized) trials and care, both in patients who improve as well in patients who deteriorate.

To investigate the accuracy of existing MCIDs in patients that are assessed for deterioration, we performed a post-hoc analysis of the immunoglobulin overtreatment in CIDP (IOC) trial, a double-blind randomized IVIg withdrawal trial investigating the need of ongoing IVIg treatment.[15] The aims of this study are to: (a) investigate the variability in scores of the I-RODS, grip strength, and MRC-SS on visits without meaningful deterioration according to a clinical anchor during observation; (b) investigate the sensitivity and specificity of existing MCIDs and combinations of MCIDs of different outcome measures for deterioration; (c) define optimized anchor-based MCIDs for deterioration for the I-RODS, grip strength, and MRC-SS.

2 | MATERIALS AND METHODS

This is a post-hoc analysis using data from the IOC trial,[15] a double-blind randomized trial in which 60 CIDP patients stable on intravenous immunoglobulin (IVIg) maintenance treatment were randomized to step-wise IVIg withdrawal (n = 29) or continuation of IVIg treatment (n = 31). The study protocol is described in full elsewhere.[15] In short, patients were assessed for deterioration during a period of 24 weeks with planned visits every 6 weeks and unplanned visits when necessary. In the IOC trial, an early endpoint was defined as decrease on the I-RODS (meaning worsening in function) of at least the MCID based on individualized standard errors (MCID-SE ≤−1.96) or if the patient or clinician judged an adverse change to be relevant. Patients who reached an early endpoint based on these predefined criteria, were unblinded and subsequently, the treating clinician decided whether restarting or increasing IVIg treatment was indicated. The Ethics Committee of Amsterdam UMC approved the study and all patients provided signed informed consent.

2.1 | Clinical data

For this study, we collected data concerning demographics, CIDP subtype, treatment allocation, and treatment details. For each visit, we collected the following outcome measures: the I-RODS, grip strength (the maximum score out of three measurements for both hands separately using a Martin-Vigorimeter), MRC-SS (ranging from 0 to 60 points), and patient's perception of clinical deterioration or improvement on a 5-point patient global impression of change scale (PGIC ranging from much better to much worse) compared with study entry. The INCAT disability scale was not used in the trial and therefore not analyzed. Treatment decision of physician (restart or increase of IVIg dose after reaching a predefined early endpoint or continuation of study treatment) was recorded for each visit.

2.2 | Minimal clinically important differences definitions

We included the most frequently used MCIDs per outcome measure in CIDP,[16] which may be distribution-based, anchor-based, or based on expert opinion. If different MCID cutoffs were reported in literature, we investigated multiple MCIDs per outcome measure to compare accuracy. For the I-RODS, we used a MCID-SE of ±1.96 that was suggested in the original paper by Draak et al.,[9] which was also one of the criteria to define an early endpoint in the IOC trial. We also included a change of ±4 centile points on the metric scale as MCID that has been used increasingly in recent trials, despite its arbitrary choice.[17,18] For grip strength, we used a change of ±8 kilopascal (kPa) and ±14 kPa as MCIDs. In most trials, a change of ±8 kPa has been used,[17,19] however due to lack of specificity of this change, more recently a larger change of ±14 kPa has been proposed.[12,13] MCID for the MRC-SS is also still debated, and most studies have suggested either a change of ±2 or ±4 points on a scale of 0 to 60 as MCIDs.[8]

2.3 | External anchors for deterioration

We used a "treatment anchor" as primary external anchor, which was based on treatment changes implemented by the treating physician after an early endpoint was reached according to the study definitions, which could be either the MCID-SE ≤−1.96 on the I-RODS or if the patient or clinician judged an adverse change to be relevant (see above). We chose this anchor assuming that a treatment change indicates that both patient and physician judge the deterioration to be clinically relevant. Visits were categorized by the treatment anchor as follows: when an early endpoint visit was followed by restart of treatment or increase of dose, the visit was regarded as a visit with meaningful deterioration according to the treatment anchor (VD + TA); when the treatment regimen remained unchanged, the visit was regarded as a visit without meaningful deterioration according to the treatment anchor (VD − TA). In addition, we used a patient anchor that was based on change on the PGIC, which was dichotomized into deterioration (slightly and much worse on the PGIC) or no deterioration (about the same and slightly and much better on the PGIC). Visits were defined as visits with meaningful deterioration according to the patient anchor (VD + PA) and visits without meaningful deterioration according to the patient anchor (VD − PA). In this study, we did not define visits as visits with improvement using these anchors, as patients were not primarily assessed for improvement during this trial.

2.4 | Analysis

We calculated change scores for all outcome measures between each follow-up and baseline visit. For grip strength, we calculated change scores for both hands and selected the hand with the maximum decrease for further analyses. For the analyses of the I-RODS, we used centile scores and the MCID-SE, which was calculated using logit-transformed scores and their corresponding individualized standard errors (SEs), based on previous Rasch analyses.[9]

We calculated the variability in scores at VD − TA, using the limits of agreement based on the Bland and Altman method.[20] We used intercept-only multilevel models to allow repeated measurements, containing the change scores as dependent variables and the patients as random effects. The following formula was used: the mean difference (the fixed intercept of the multilevel model) ± 1.96 * SD (√[the sum of the within-patient and between-patient variance]). In other words, we calculated the 95% confidence interval (CI) of change scores that were observed on visits without a change in treatment regimen because of deterioration (VD − TA). These limits of agreement represent the limits of the variability; changes within these limits are not significantly different from changes observed in patients without a meaningful deterioration.

To investigate diagnostic accuracy, we first determined the discriminatory potential of the I-RODS, grip strength and MRC-SS by creating receiver operating characteristic (ROC) curves using cross-sectional data of the last visit of the randomized controlled trial (RCT). Next, we calculated sensitivities and specificities of existing MCIDs using both anchors as reference standards.

In addition, we defined MCIDs optimized for deterioration based on our dataset, by selecting two different cutoffs in the ROC curves: based on the Youden's index and based on the maximized specificity with a minimal sensitivity of 60%. The Youden's index gives an indication of the performance of a diagnostic test by combining the values of the sensitivity and specificity using the following formula: sensitivity + specificity − 1, which ranges from 0 (worst) to 1 (best).[21] We favored specificity over sensitivity because erroneously diagnosing deterioration has more implications in a clinical setting, as this would imply restarting treatment for a long period of time whereas, if patients do not meet the criteria for deterioration, careful follow-up of patients can adequately reveal a missed deterioration.

We also explored the accuracy of combining existing and optimized MCIDs of various outcome measures on the level of disability (using only centile score of the I-RODS due to the extensiveness of analyses) and impairment by using an "and" and/or "or" function. For example, for the MCID of "−4 centile on the I-RODS and −2 points on the MRC-SS", patients had to decrease at least 4 centiles on the I-RODS *as well as* at least 2 points on the MRC-SS to be classified as deteriorated by the combined MCID. This means that a patient that deteriorated at least the MCID of −4 centiles on the I-RODS, but did not deteriorate at least the MCID of −2 points on the MRC-SS, was not classified as deteriorated by this combined MCID for deterioration. For the combined MCID of "−4 centile on the I-RODS or −2 points on the MRC-SS", patients had to decrease at least 4 centiles *or* at least 2 points on the MRC-SS to be classified as deteriorated by the combined MCID. For the combination of the "I-RODS and either grip strength or MRC-SS", patients had to decrease at least the MCID on the I-RODS *and* had to decrease at least the MCID on *either* grip strength *or* the MRC-SS. We selected the most accurate combined MCIDs for the tables and figures in the primary article. This selection of combined MCIDs was based on a minimal sensitivity and specificity of 60%, and at least a Youden's index of 0.5.

We plotted the sensitivity and specificity of all existing and new, optimized MCIDs and a selection of most accurate combined MCIDs.

Analyses were performed in R (version 4.0.2.),[22] a software environment for statistical computing and graphics.

3 | RESULTS

Sixty patients were included in the study, of whom a total of 241 visits were available for analysis. The baseline characteristics of all patients can be found in Table 1. In 24/60 (40%) patients, meaningful deterioration was noted during the study follow-up of 24 weeks as based on the primary treatment anchor, while 28/56 (50%) had deteriorated according to the secondary patient anchor. Twenty-nine patients were allocated to IVIg withdrawal, of whom 17 reached an early endpoint (59%), which resulted in restarting IVIg treatment or increasing its dosage in all (17/17) patients. Thirty-one patients were allocated

TABLE 1 Baseline demographic and clinical characteristics at baseline (n = 60 patients)

	Patients with meaningful deterioration during trial according to treatment anchor (n = 24)	Patients without meaningful deterioration during trial according to treatment anchor (n = 36)
Sex (n males [%])	17/24 (71%)	25/36 (69%)
Age (mean [SD; range])	61 years (SD 14, range 30-86)	57 years (SD 15, range 21-81)
CIDP phenotype		
Typical	20/24 (83%)	27/36 (75%)
Atypical	4/24 (17%)	9/36 (25%)
Asymmetric CIDP	2/24 (8%)	4/36 (11%)
Pure motor/sensory	2/24 (8%)	5/36 (14%)
Treatment allocation during trial		
IVIg continuation	7/24 (29%)	24/36 (67%)
IVIg withdrawal	17/24 (71%)	12/36 (33%)
Disease duration (median [range])	64 months (17-586)	49 months (7-314)
Wear off symptoms	8/24 (33%)	7/36 (19%)
Duration of IVIg treatment		
6-12 months	9/24 (38%)	19/36 (53%)
>12 months	15/24 (63%)	17/36 (47%)
Previous withdrawal attempts	12/24 (50%)	11/36 (31%)
IVIg interval (median [range])	3 weeks (2-6)	3 weeks (2-6)
IVIg dose per infusion (median [range])	40 g (10-80)	50 g (30-80)
I-RODS (median [IQR, range])		
Logits	3.09 (IQR 1.47-6.25, range −2.46 to 8.11)	5.03 (IQR 3.66-6.25, range −0.61 to 8.11)
Centiles	67 (IQR 56-88, range 30-100)	78 (IQR 71-88, range 42-100)
MRC-SS (median [IQR; range])	59 (IQR 56-60, range 38-60)	60 (IQR 58-60, range 51-60)
Grip strength, right hand (mean [SD; range])	77 kPa (SD 39; 9-155)	83 kPa (SD 25; 42-135)
Grip strength, left hand (mean [SD; range])	69 kPa (SD 38; 0-155)	82 kPa (SD 27; 40-135)

Note: Grip strength was measured using a Martin-Vigorimeter ranging from 0 to 160 kPa. The MRC-SS used for analyses ranged from 0 (total paralysis) to 60 points (normal strength).
Abbreviations: I-RODS, Inflammatory Rasch-built Overall Disability Scale; IQR, interquartile range; IVIg, intravenous immunoglobulins; kPa, kilopascal; MRC-SS, MRC sum score; SD, standard deviation.

to IVIg continuation of whom 13 (42%) reached an early endpoint, which resulted in increasing their IVIg dosage in 7/13 (54%) patients. The reason for an early endpoint was deterioration based on MCID-SE in 50% (15/30; 5/13 IVIg continuation and 10/17 IVIg withdrawal) of patients, clinician-based in 23% (7/30; 3/13 IVIg continuation and 4/17 IVIg withdrawal) or patient-based in 27% (8/30; 5/13 IVIg continuation and 3/17 IVIg withdrawal).

3.1 | Anchor-based change scores

The means and distributions of change scores for I-RODS, grip strength and MRC-SS based on the primary anchor for meaningful deterioration (VD + TA and VD − TA) can be found in Table 2 and corresponding Figure 1. In addition, the means and distributions of change scores of the I-RODS, grip strength and MRC-SS categorized based on the secondary anchor (VD + PA and VD − PA) can be found in Table A1. Twenty-four visits (13%) were classified as visits with meaningful deterioration when using the treatment anchor (VD + TA), while 54 visits (29%) were classified as visits with meaningful deterioration when applying the patient anchor (VD + PA). Changes in outcome measures on VD + TA generally appeared larger compared with VD + PA. Together, this suggest that deterioration as perceived by patients did not always lead to change in treatment.

3.2 | Visits without deterioration

The limits of agreement, representing the limits of variability, for the I-RODS, grip strength, and MRC-SS can be found in Table 3. These limits are based on the 95% CI of change scores that were observed

TABLE 2 Change scores of visits with and without meaningful deterioration according to the treatment anchor for the I-RODS, grip strength and MRC-SS (VD + TA and VD − TA, n = 181 visits of 60 patients)

Category of visit according to treatment anchor	n visits	I-RODS, centiles (mean change [SD])	I-RODS, MCID-SE (mean change [SD])	Grip strength, kPa (mean change [SD])	MRC-SS, points (median change [IQR])
Visits without meaningful deterioration (VD − TA)	157 visits of 52 patients	0 (SD 8)	−0.05 (SD 0.97)	−4 (SD 14)	0 (IQR 0–0)
Visits with meaningful deterioration (VD + TA)	24 visits of 24 patients	−11 (SD 8)	−1.95 (SD 1.30)	−18 (SD 19)	−0.5 (IQR −4 to 0)

Note: For grip strength, change scores of the hand with the maximum decrease are shown measured using a Martin-Vigorimeter ranging from 0 to 160 kPa. The MRC-SS used for analyses ranged from 0 (total paralysis) to 60 (normal strength).
Abbreviations: I-RODS, Inflammatory Rasch-built Overall Disability Scale; IQR, interquartile range; kPa, kilopascal; MCID, minimal clinically important differences; MCID-SE, individual change score divided by their corresponding stand error of the difference; MRC-SS, MRC sum score; SD, standard deviation; VD + TA, visit with meaningful deterioration according to the treatment anchor; VD – TA, visit without meaningful deterioration according to the treatment anchor.

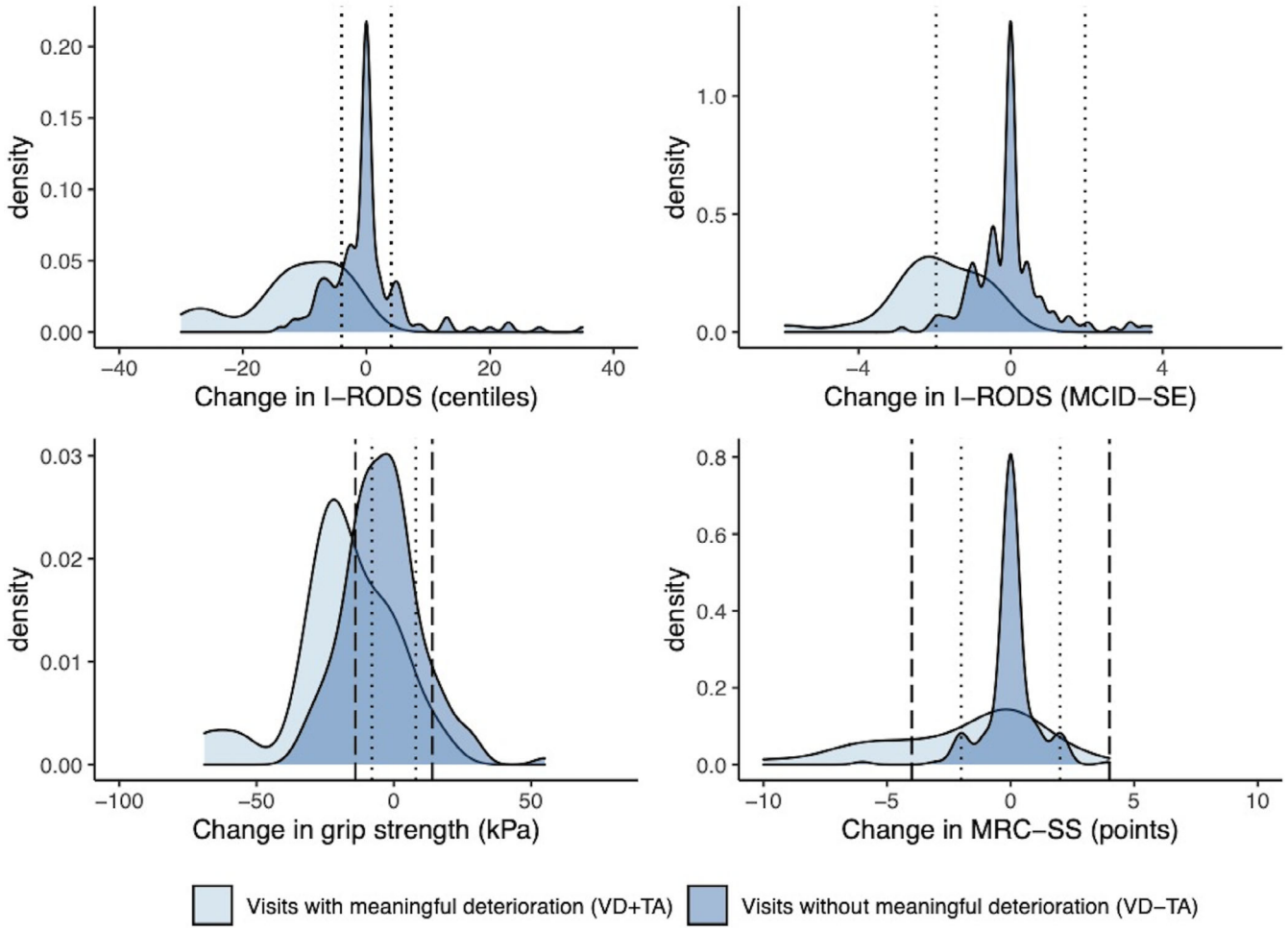

FIGURE 1 Change scores of visits with and without meaningful deterioration according to the treatment anchor for the I-RODS, grip strength and MRC-SS (VD + TA and VD − TA, n = 181 visits of 60 patients). The dotted and striped lines represent the current MCIDs. The dotted lines represent the following MCIDs: ±4 centiles on the I-RODS, ±1.96 MCID-SE on the I-RODS, ±8 kPa for grip strength, and ±2 points on the MRC-SS. The striped lines represent the MCIDs of ±14 kPa on grip strength, and ±4 points on the MRC-SS. For grip strength, change scores of the hand with the maximum decrease are shown measured using a Martin-Vigorimeter ranging from 0 to 160 kPa. The MRC-SS used for analyses ranged from 0 (total paralysis) to 60 (normal strength). I-RODS, Inflammatory Rasch-built Overall Disability Scale; kPa: kilopascal; MCID, minimal clinically important differences; MCID-SE, individual change score divided by their corresponding stand error of the difference; MRC-SS, MRC sum score; VD + TA, visit with meaningful deterioration according to the treatment anchor; VD − TA, visit without meaningful deterioration according to the treatment anchor

on VD − TA. The limits of agreement of the I-RODS ranged between −14 to 12 centiles and −2.07 to 1.80 MCID-SE. For grip strength, the limits of agreement ranged from −33 to 25 kPa and for MRC-SS from −2.2 to 2.0 points. In other words, these limits represent the 95% CIs of change scores that were observed on visits without a treatment change due to deterioration.

Deterioration in outcome measures exceeding an MCID on VD-TA varied between 1% and 42%, depending on the outcome measure and cutoff (Table 4). The MCID cutoff of 8 kPa on grip strength was considered the least specific as it was observed in 66/157 (42%) of visits without leading to a treatment change. That means that on those visits, a meaningful deterioration on the outcome measure was established when applying the MCID definitions, but the treatment was not increased or restarted. Similar percentages of changes in outcome measures exceeding an MCID for deterioration (between 0% and 40% of visits) were observed in

TABLE 3 Limits of agreement for the I-RODS, grip strength, and MRC-SS

Outcome measure	Mean difference (95% CI)	Lower limit of agreement	Upper limit of agreement
I-RODS, centiles	−1 (−2, 1)	−14	12
I-RODS, MCID-SE	−0.14 (−0.39, 0.11)	−2.07	1.80
Grip strength, kPa	−4 (−8, 0)	−33	25
MRC-SS, points	−0.1 (−0.3, 0.2)	−2.2	2.0

Note: Calculations are based on change scores on 157 visits without meaningful deterioration based on the treatment anchor (VD − TA) of 52 patients using intercept-only multilevel models to allow repeated measurements with the following formula: the mean difference (the fixed intercept of the multilevel model) ± 1.96 * SD (√ [the sum of the within-patient and between-patient variance]). These limits of agreement represent the limits of the variability; changes within these limits are not significantly different from changes observed in patients without meaningful deterioration.
For grip strength, change scores of the hand with the maximum decrease are shown measured using a Martin-Vigorimeter ranging from 0 to 160 kPa. The MRC-SS used for analyses ranged from 0 (total paralysis) to 60 (normal strength).
Abbreviations: I-RODS, Inflammatory Rasch-built Overall Disability Scale; IQR, interquartile range; kPa, kilopascal; MCID, minimal clinically important differences; MCID-SE, individual change score divided by their corresponding stand error of the difference; MRC-SS, MRC sum score; SD, standard deviation.

TABLE 4 MCIDs during visits *without meaningful* deterioration according to the treatment anchor (VD-TA, n = 157 visits of 52 patients)

Outcome measure	MCID cutoff for deterioration	No MCID for deterioration (%; n/N visits of n patients)	MCID deterioration (%; n/N visits of n patients)
Singular MCIDs			
I-RODS	−4 centiles	77% (122/157 of 42 patients)	23% (35/157 visits of 22 patients)
	−1.96 MCID-SE	98% (154/157 of 51 patients)	2% (3/157 of 3 patients)
Grip strength	−8 kPa	58% (91/157 of 36 patients)	42% (66/157 of 30 patients)
	−14 kPa	76% (119/157 of 45 patients)	24% (38/157 of 20 patients)
MRC-SS	−2 points	92% (144/157 of 48 patients)	8% (13/157 of 11 patients)
	−4 points	99% (156/157 of 52 patients)	1% (1/157 of 1 patient)
Combined MCIDs			
I-RODS and either grip strength or MRC-SS	−4 centiles and either −8 kPa or −2 points	92% (144/157 visits of 48 patients)	8% (13/157 of 10 patients)
	−4 centiles and either −14 kPa or −2 points	96% (150/157 of 49 patients)	4% (7/157 of 6 patients)
I-RODS or MRC-SS	−4 centiles or −2 points	71% (112/157 of 40 patients)	29% (45/157 of 30 patients)
	−1.96 MCID-SE or −2 points	90% (142/157 of 42 patients)	10% (15/157 of 13 patients)

Note: The table shows the percentages of visits without meaningful deterioration (VD − TA) during which any of the existing MCIDs were reached using change scores between each follow-up visit and the baseline visit. No MCID for deterioration was reached when the change score was >—the MCID value and the MCID for deterioration was reached when the change score was ≤—the MCID value.
For the combined MCID of "−4 centile on the I-RODS or −2 points on the MRC-SS", patients had to decrease at least 4 centiles *or* at least 2 points on the MRC-SS to be classified as deteriorated by the combined MCID. For the combined MCID of the "I-RODS and either grip strength or MRC-SS", patients had to decrease at least the MCID on the I-RODS *and* had to decrease at least the MCID on *either* grip strength *or* the MRC-SS.
For grip strength, change scores of the hand with the maximum decrease are shown measured using a Martin-Vigorimeter ranging from 0 to 160 kPa. The MRC-SS used for analyses ranged from 0 (total paralysis) to 60 (normal strength).
Abbreviations: I-RODS, Inflammatory Rasch-built Overall Disability Scale; kPa, kilopascal; MCID, minimal clinically important differences; MCID-SE, individual change score divided by their corresponding stand error of the difference; MRC-SS, MRC sum score.

the VD-PA (Table A2). This means that on those visits, a meaningful deterioration on the outcome measure was established when applying the MCID definitions, but patients reported not to feel deteriorated.

The percentages of visits with deteriorations exceeding the most accurate combined MCIDs on VD − TA and VD − PA, can also be found in Table 4 and Table A2, respectively.

Finally, we investigated discrepancies between the treatment and patient anchor. Of all VD − TA, 20% (31/154) were classified as VD + PA, which means that although patients reported to feel deteriorated, the treatment was not always restarted or increased.

3.3 | Visits with deterioration

During VD + TA, all patients reported to feel deteriorated on the PGIC except one patient that felt improved (95%, 21/22). The sensitivities, specificities and Youden's indexes of all existing and optimized MCIDs are shown in Figure 2 and Table 5. The MCID-SE (I-RODS) and MRC-SS MCIDs showed highest specificities. Sensitivities for existing MCIDs were suboptimal except for the MCID of −4 centiles on the I-RODS at the expense of specificity when using the treatment anchor as reference. Generally, the optimized MCID cutoffs for deterioration differed based on the method used for calculations, that is,

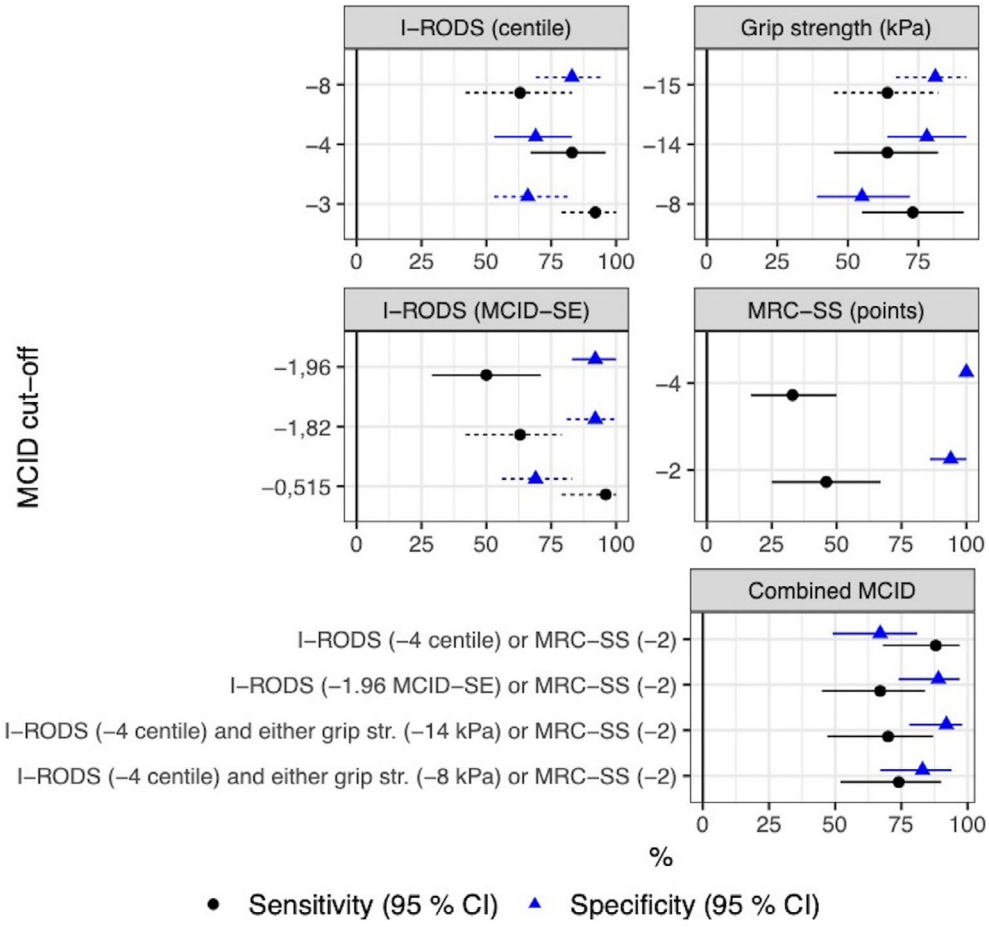

FIGURE 2 Sensitivities and specificities of existing MCIDs (full line), combinations of existing MCIDs, and MCIDs optimized (dotted) for deterioration of the I-RODS, grip strength and MRC-SS. The MCID values are shown on the y-axis, which are either existing (full line) or calculated based on our sample (dotted line). The symbol represents either the sensitivity or specificity on the x-axis and the line the 95% confidence intervals, based on the treatment anchor. The sensitivity constitutes of the percentage of patients in which meaningful deterioration is correctly identified by the MCID of all patients with meaningful deterioration. The specificity constitutes of the percentage of the patients in which the absence of meaningful deterioration is correctly identified by the MCID of all patients without meaningful deterioration. The confidence intervals of the sensitivity and specificity are based on 2000 stratified bootstrap replicates. Combinations between outcome measures were made by using an "and" and/or "or" function. For the combined MCID of "−4 centile on the I-RODS or −2 points on the MRC-SS", patients had to decrease at least 4 centiles *or* at least 2 points on the MRC-SS to be classified as deteriorated by the combined MCID. For the combination of the "I-RODS and either grip strength or MRC-SS", patients had to decrease at least the MCID on the I-RODS *and* had to decrease at least the MCID on *either* grip strength *or* the MRC-SS. In Table A3, the data are shown in full detail. I-RODS, Inflammatory Rasch-built Overall Disability Scale; kPa, kilopascal; MCID, minimal clinically important differences; MCID-SE, individual change score divided by their corresponding stand error of the difference; MRC-SS, MRC sum score

TABLE 5 Accuracy of existing MCIDs and MCIDs optimized for deterioration of the I-RODS, grip strength and MRC-SS

Outcome measure	MCID calculation method	MCID cutoff for deterioration	Sensitivity (95% CI)	Specificity (95% CI)	Youden's index
I-RODS	Optimized, Youden's index	−3 centiles	92% (79-100)	66% (50-81)	0.58
	Existing MCID	−4 centiles	83% (67-96)	69% (53-83)	0.52
	Optimized, max. Specificity	−8 centiles	63% (42-83)	83% (69-94)	0.46
	Optimized, Youden's index	−0.515 MCID-SE	96% (79-100)	69% (56-83)	0.65
	Optimized, max. Specificity	−1.82 MCID-SE	63% (42-79)	92% (81-100)	0.55
	Existing MCID	−1.96 MCID-SE	50% (29-71)	92% (83-100)	0.42
Grip strength	Existing MCID	−8 kPa	73% (55-91)	55% (39-72)	0.28
	Existing MCID	−14 kPa	64% (45-82)	78% (64-92)	0.42
	Optimized, Youden's index and optimized, max. Specificity	−15 kPa	64% (45-82)	81% (67-92)	0.45
MRC-SS	Optimized, max. Specificity	NA[a]	-	-	-
	Existing MCID and Optimized, Youden's index	−2 points	46% (25-67)	94% (86-100)	0.40
	Existing MCID	−4 points	33% (17-50)	100% (100-100)	0.33

Note: The sensitivities and specificities were calculated using the treatment anchor, based on data of the last visit of the RCT. The Youden's index combines the values of the sensitivity and specificity using the following formula: sensitivity + specificity − 1, and ranges from 0 (worst) to 1 (best). For calculation of the MCID based on the maximized specificity, the cutoff with the highest specificity was selected with a minimum sensitivity of 60%.
Abbreviations: CI, confidence intervals; I-RODS, Inflammatory Rasch-built Overall Disability Scale; kPa, kilopascal; max., maximized; MCID, minimal clinically important differences; MCID-SE, individual change score of the I-RODS divided by their corresponding stand error of the difference; MRC-SS, MRC sum score.
[a]Not applicable; the lower limit of 60% sensitivity was not associated with deterioration. The cutoff had therefore no clinical meaning and was not shown.

either the Youden's index or based on the maximized specificity. In addition, the sensitivities and specificities based on the patient anchor can be found in Table A3.

The sensitivities and specificities of all the combined MCIDs based on our data can be found in Table A4. The highest specificity, of 100% (95% CI: 90-100), was found when applying the combined MCID that encompassed both −1.96 SE-MCID on the I-RODS *and* −14 kPa of grip strength, with a sensitivity of 35% (95% CI: 16-57). The combined MCID of −3 centiles on the I-RODS *or* −15 kPa of grip strength (both optimized MCIDs), resulted in the highest sensitivity of 100% (95% CI: 86-100) with a specificity of 53% (95% CI: 35-70). The combined MCID based on existing MCIDs with the highest sensitivity encompassed −4 centile on the I-RODS *or* −8 kPa of grip strength, which had a sensitivity of 96% (95% CI: 79-100) with a specificity of 42% (95% CI: 26-59). We found similar accuracies of the combinations of existing MCIDs and the combinations of optimized MCIDs. The sensitivities and specificities of the selection of the most accurate combined MCIDs—the selection was based on a minimal sensitivity and specificity of 60%, and Youden's index of at least 0.5—can be found in Figure 2. When applying a minimal sensitivity of 60%, we found that the combined MCID of the I-RODS (−4 centiles) with either grip strength (−14 kPa) or MRC-SS (−2 points), resulted in the highest specificity of 92% (95% CI: 78-98) with sensitivity of 70% (95% CI: 47-87) and the highest Youden's index. This combined MCID was identical in terms of sensitivity, specificity, and Youden's index to the combined MCID that encompassed optimized MCIDs of the I-RODS (−3 centiles) with either grip strength (−15 kPa) or MRC-SS (−2 points), which is therefore not shown in Figure 2, Table 2, and Table A4.

Most MCIDs or combinations of MCIDs had large CIs of the sensitivity and specificity. Only one MCID cutoff, either existing or optimized based on our data, was outside the limits of the variability observed in VD − TA (ie, the MRC-SS >4 MCID). In other words, most MCID cutoffs were within the CI of changes that were observed on visits without a treatment change due to deterioration.

4 | DISCUSSION

In this study, we found substantial variability in scores of the I-RODS, grip strength and MRC-SS in visits without meaningful deterioration when measured in patients participating in a randomized trial. Most existing MCIDs used for deterioration were within the limits of this variability. Moreover, we found that both existing MCIDs and the optimized anchor-based MCIDs from this study were insensitive but highly specific for detecting deterioration or vice versa, limiting their use on the individual level.

The considerable variability in outcome measures on visits without meaningful deterioration according to the treatment or patient anchor we observed may be explained by various factors, including minor (treatment related) symptom fluctuations, changes in mood, pain or fatigue and inconsistencies due to misunderstanding of the items or answer categories of questionnaire, poor interrater reliability for the MRC-SS, or suboptimal method of measurement such when

assessing grip strength.[10,12,16,23-25] In addition, specifically applicable for our study, variability may be also influenced by blinding of patients and potential nocebo effects.[15] Recently, a retrospective study that included stable, IgG-treated CIDP patients found that, during routine care, the MCIDs of the 4 centiles of I-RODS, 8 kg of grip strength (using a Jamar dynamometer) and 2 points on the MRC-SS were met in 44%, 11% and 44%, respectively.[26] This implies that variability exceeding MCIDs in stable patients is also common in clinical care. Another study of CIDP patients investigated the minimum detectable change (MDC),[27] which is conceptually similar to the limits of variability used in our study, as both reflect the threshold of changes beyond that what can be observed in stable patients. The MDC of I-RODS was smaller than limits of variability in our data.[27] Also, contrary to our results, the distribution-based MCID value of grip strength and the I-RODS exceeded the MDC,[27] indicating that clinically meaningful changes can be distinguished from random variability. The inconsistency with our results can be explained by the authors' use of reliability parameters used for calculations of the MDC, which were derived from studies containing mainly patients with stable Guillain-Barré syndrome, while stable CIDP patients only accounted for 27% (80/294) and 29% (30/102) of patients, in studies of the I-RODS and grip strength respectively.[14,28] Variability in stable patients might be disease specific, especially since stable CIDP patients often experience (treatment related) fluctuations in symptoms and patients' reported disability is subjected to variables such as fatigue and mood.[12,13,23,24]

To date, no studies have investigated accuracies of MCIDs for deterioration using a clinical anchor as reference standard. One study reported that the MCID of −8 of grip strength was more sensitive in detecting meaningful deterioration than the MCID of 1 point of the INCAT-DS in the placebo group of the ICE trial.[19] Another study reported insufficient sensitivity to detect meaningful deterioration using the MCID of 1 point on the INCAT-DS and changed the protocol amendment by adding the MCIDs of −8 kPa of grip strength and 4 centile points on the I-RODS to increase sensitivity.[29]

We found some inconsistency between the patient anchor and the treatment anchor for deterioration. On some visits, patients reported to feel deteriorated, but the treatment was not increased or restarted on that visit. It is possible that the physician attributed the patient reported deterioration to symptoms unrelated to active CIDP or that the deterioration experienced by patients could not be objectified by the clinician. Indeed, we found that changes in outcome measures tended to be smaller when deterioration was classified according to the patient compared with according to the clinician. Differences in perspectives of change between patients and physicians have been reported before.[30] Currently, no consensus has been reached regarding the best anchor to define a meaningful change.[30]

As expected, the MCID combinations resulted in various sensitivities and specificities. The combination of the I-RODS of −4 centiles and either grip strength or MRC-SS resulted in high specificities as well as moderate sensitivities, which may be preferable for detecting meaningful deterioration to avoid overtreatment. We did not perform analysis on whether our patients had isolated distal or proximal weakness. Nonetheless, the high accuracy of this combined MCID implicates that meaningful deterioration is not always accompanied by change in both grip strength and MRC-SS. However, these results need to be validated in larger studies. Combination of the optimized MCIDs did not result in better diagnostic accuracy than the combinations of currently used MCIDs.

The main limitation of our study was the sample size. Although most results were based on multiple measurements per patient, some results were based on a single measurement per patient and showed wide CIs. Moreover, the sample size precluded investigation of additional determinants such as disease phenotype and baseline severity of disability or impairment on the accuracies of the MCIDs. It is likely that the cutoff for a meaningful change differs between a severely and slightly disabled patient, although most patients in the IOC trial had limited disability. In addition, we chose a treatment decision as primary anchor rather than a patient anchor, because patient anchors are often criticized for recall bias and questionable reliability.[31] However, it is unclear which anchor best represents a meaningful change. Importantly, we used data from an RCT in which patients were assessed regularly. We cannot exclude the possibility that patients enrolled in an RCT answer patient-reported outcome measures differently, for example, due to expectations of the allocated treatment,[32] compared with patients in regular care and therefore our results might not be generalizable to normal care. In addition, I-RODS changes were used to define an early endpoint, although this endpoint was only reached by half of patients and decision to restart treatment after an early endpoint was made by the treating physician, we cannot rule out that this decision to treat was influenced by the I-RODS change.

4.1 | Potential implications for clinicians

Our results highlight the uncertainties concerning interpreting MCIDs in research and probably clinical care. In general, and specifically for grip strength, which is known for daily fluctuations, it is advised to use multiple consecutive time points to improve the distinction between meaningful changes and random variability, by using the means of multiple time points or by using a minimum number of days with consistent change.[12,23] However, the minimal time with consistent change to be considered clinically meaningful has not been investigated. To make further recommendations concerning MCIDs, additional research is needed to investigate the variability of scores in stable patients and investigate the accuracies and cutoffs of (combined) MCIDs for both improvement and deterioration, using larger sample sizes, in the context of both (randomized) clinical trials and clinical care, for example by means of a meta-analysis. It is advised to combine MCIDs on the level of disability as well as impairment to assess changes.[26,33] The MCID calculations should be based primarily on relevant and multiple clinical anchors and methods, since it is unknown which anchors and methods are superior. Distribution-based methods should be used alongside clinical

anchors to define the threshold beyond the random variability. Finally, a single objective "optimal" MCID cutoff will most likely not exist, since different contexts warrant a different sensitivity-specificity balance. In case of deterioration, a higher specificity combined with frequent monitoring might be appropriate to avoid over-treatment in CIDP. Ideally, the final selection of MCIDs in various contexts should be based on extensive review and group consensus among clinical experts and patient groups.

ACKNOWLEDGEMENTS

We thank Janneke Zwiers and Mary Muijs from Sanquin plasma products B.V., the research nurses from Penthecilia B.V. and Margit Dros from Mediq Tefa for ensuring the blinding and distribution of the IMP. We also like to thank Michael Lunn, Joke Dijk and Brent Opmeer for participating in the DSMB.

FUNDING INFORMATION

Study was made possible by grants from ZonMw (836021010) and Prinses Beatrix Spierfonds (Dutch Charity for neuromuscular diseases). *Sanquin Plasma Products B.V.* provided the placebo, preparation, blinding and distribution of the study treatment. The funders had no role in the trial design, data collection, data analysis, data interpretation, or the writing of the report.

CONFLICT OF INTEREST

F. Eftimov report grants from ZonMw (Dutch governmental agency) for the IOC trial in which data was collected for this study, and a grant from Prinses Beatrix Spierfonds (Dutch charity) on clinimetrics in CIDP. Outside of the submitted work, he has received other grants from ZonMw and Prinses Beatrix Spierfonds for studies in CIDP. As principal investigator of INCbase, the also reports investigator-initiated grants from Kedrion, Terumo BCT, CSL-Behring, Grifols and Takeda Pharmaceutical Company. His institution has received fees from UCB Pharma, CSL Behring, Grifols and Takeda for advisory board membership and/or lectures. All grants and fees were paid to his institution. He is a member of the Cochrane Neuromuscular Editorial Board. L. Wieske received research grants from Grifols (2019) and the GBS/ CIDP Foundation (2020) outside the submitted word. I. van Schaik reports grants from Dutch Governmental grant (ZonMw/Rational Pharmacotherapy program, non-financial support from Sanquin Plasma Products B.V., during the conduct of the study; other from CSL-Behring, outside the submitted work. I. S. J. Merkies reports grants from GBS/CIDP Foundation International, grants from Talecris Talents program foundation, reports personal fees and other from Steering committee member of several studies: ICE trial, CSL Behring, LFB, Novartis, Baxter, UCB, Octapharma, Argenx, outside the submitted work. M. E. Adrichem, I. Lucke, and R. van Veen have nothing to disclose.

DATA AVAILABILITY STATEMENT

The corresponding author (F.E.) has full access to the IOC study protocol and all the data in the study. Data are available upon reasonable request.

ORCID

Robin van Veen https://orcid.org/0000-0001-7702-4930

REFERENCES

1. Allen JA, Merkies ISJ, Lewis RA. Monitoring clinical course and treatment response in chronic inflammatory demyelinating polyneuropathy during routine care. *JAMA Neurol*. 2020;77:1159-1166. doi:10.1001/jamaneurol.2020.0781
2. Vanhoutte EK, Faber CG, Merkies IS, PeriNomS Study Group. 196th ENMC international workshop: outcome measures in inflammatory peripheral neuropathies 8-10 February 2013, Naarden, The Netherlands. *Neuromuscul Disord*. 2013;23(11):924-933. doi:10.1016/j.nmd.2013.06.006 (published Online First: 2013/07/10).
3. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407-415. doi:10.1016/0197-2456(89)90005-6
4. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102-109. doi:10.1016/j.jclinepi.2007.03.012 (published Online First: 2008/01/08).
5. Draak THP, de Greef BTA, Faber CG, Merkies ISJ, The PeriNomS Study Group. The minimum clinically important difference: which direction to take. *Eur J Neurol*. 2019;26(6):850-855. doi:10.1111/ene.13941 (published Online First: 2019/02/23).
6. de Vet HC, Terwee CB. The minimal detectable change should not replace the minimal important difference. *J Clin Epidemiol*. 2010;63(7):804-805; author reply 806. doi:10.1016/j.jclinepi.2009.12.015 (published Online First: April 20, 2010).
7. Wright A, Hannon J, Hegedus EJ, Kavchak AE. Clinimetrics corner: a closer look at the minimal clinically important difference (MCID). *J Man Manip Ther*. 2012;20(3):160-166. doi:10.1179/2042618612Y.0000000001 (published Online First: 2013/08/02).
8. Merkies IS, van Nes SI, Hanna K, et al. Confirming the efficacy of intravenous immunoglobulin in CIDP through minimum clinically important differences: shifting from statistical significance to clinical relevance. *J Neurol Neurosurg Psychiatry*. 2010;81(11):1194-1199. doi:10.1136/jnnp.2009.194324 (published Online First: 2010/07/22).
9. Draak TH, Vanhoutte EK, van Nes SI, et al. Changing outcome in inflammatory neuropathies. *Neurology*. 2014;83(23):2124-2132. doi:10.1212/WNL.0000000000001044
10. Vanhoutte EK, Draak TH, 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
11. Rajabally YA, Ghasemi M. Comparative value and determinants of suitability of outcome measures in treated chronic inflammatory demyelinating polyneuropathy. *Muscle Nerve*. 2020;61(2):182-186. doi:10.1002/mus.26747 (published Online First: 2019/10/28).
12. Doneddu PE, Hadden RDM. Daily grip strength response to IVIg in chronic immune neuropathies. *Muscle Nerve*. 2020;62:103-110. doi:10.1002/mus.26898 (published Online First: 2020/04/23).
13. Doneddu PE, Mandia D, Gentile F, et al. Home monitoring of maintenance intravenous immunoglobulin therapy in patients with chronic inflammatory neuropathy. *J Peripher Nerv Syst*. 2020;25(3):238-246. doi:10.1111/jns.12396 (published Online First: 2020/05/30).
14. Draak TH, Pruppers MHJ, van Nes SI, et al. Grip strength comparison in immune-mediated neuropathies Vigorimeter vs Jamar. *J Peripher Nerv Syst*. 2015;20:269-276. doi:10.1111/jns.12126
15. Adrichem ME, Lucke IM, Vrancken AFJ, et al. Withdrawal of intravenous immunoglobulin in chronic inflammatory demyelinating polyradiculoneuropathy. *Brain*. 2022;awac054. doi:10.1093/brain/awac054
16. Allen JA, Eftimov F, Querol L. Outcome measures and biomarkers in chronic inflammatory demyelinating polyradiculoneuropathy: from research to clinical practice. *Expert Rev Neurother*. 2021;21(7):805-816. doi:10.1080/14737175.2021.1944104 (published Online First: 2021/06/17).

17. van Schaik IN, Bril V, van Geloven N, et al. Subcutaneous immunoglobulin for maintenance treatment in chronic inflammatory demyelinating polyneuropathy (PATH): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Neurol*. 2018;17(1):35-46. doi:10.1016/s1474-4422(17)30378-2
18. National Library of Medicine. Rituximab in chronic inflammatory demyelinating polyneuropathy. Clinicaltrials.gov [Internet]. 2000 Feb 29. Identifier: NCT04480450. July 21, 2020. Bethesda (MD): National Library of Medicine (US). Accessed May 12, 2022. https://clinicaltrials.gov/ct2/show/NCT04480450
19. Vanhoutte EK, Latov N, Deng C, et al. Vigorimeter grip strength in CIDP: a responsive tool that rapidly measures the effect of IVIG--the ICE study. *Eur J Neurol*. 2013;20(5):748-755. doi:10.1111/j.1468-1331.2012.03851.x (published Online First: 2012/08/16).
20. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;8(8476):307-310.
21. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32-35. doi:10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3
22. R Foundation for Statistical Computing. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/ [program]. 2014.
23. Allen JA, Pasnoor M, Dimachkie MM, et al. Quantifying treatment-related fluctuations in CIDP: results of the GRIPPER study. *Neurology*. 2021;96(14):e1876-e1886. doi:10.1212/WNL.0000000000011703 (published Online First: 2021/02/18).
24. Peric S, Bozovic I, Pruppers MHJ, et al. Validation of the Serbian version of inflammatory Rasch-built overall disability scale in patients with chronic inflammatory demyelinating polyradiculoneuropathy. *J Peripher Nerv Syst*. 2019;24(3):260-267. doi:10.1111/jns.12343 (published Online First: 2019/08/10).
25. White D, Englezou C, Rajabally YA. Patient perceptions of outcome measures in chronic inflammatory demyelinating polyneuropathy: a study of the inflammatory Rasch-built overall disability scale. *Eur J Neurol*. 2021;28(8):2596-2602. doi:10.1111/ene.14907 (published Online First: 2021/05/11).
26. Keh RYS, Selby DA, Jones S, et al. Predicting long-term trends in inflammatory neuropathy outcome measures using latent class modelling. *J Peripher Nerv Syst*. 2021;27:84-93. doi:10.1111/jns.12481
27. Rajabally YA, Afzal S, Ghasemi M. Minimal important differences and self-identifying treatment response in chronic inflammatory demyelinating polyneuropathy. *Muscle Nerve*. 2021;64(1):37-42. doi:10.1002/mus.27250 (published Online First: 2021/04/10).
28. 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.
29. Mielke O, Bril V, Cornblath DR, et al. Restabilization treatment after intravenous immunoglobulin withdrawal in chronic inflammatory demyelinating polyneuropathy: results from the pre-randomization phase of the polyneuropathy and treatment with Hizentra study. *J Peripher Nerv Syst*. 2019;24(1):72-79. doi:10.1111/jns.12303 (published Online First: 2019/01/24).
30. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol*. 2002;14(2):109-114. doi:10.1097/00002281-200203000-00006
31. Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol*. 1997;80(8):869-879. doi:10.1016/s0895-4356(97)00097-8
32. Lewis RA, Cornblath DR, Hartung HP, et al. Placebo effect in chronic inflammatory demyelinating polyneuropathy: the PATH study and a systematic review. *J Peripher Nerv Syst*. 2020;25(3):230-237. doi:10.1111/jns.12402 (published Online First: 2020/07/07).
33. Van den Bergh PYK, van Doorn PA, Hadden RDM, et al. European academy of neurology/peripheral nerve society guideline on diagnosis and treatment of chronic inflammatory demyelinating polyradiculoneuropathy: report of a joint task force-second revision. *J Peripher Nerv Syst*. 2021;26(3):242-268. doi:10.1111/jns.12455 (published Online First: 2021/06/05).

How to cite this article: van Veen R, Wieske L, Lucke I, et al. Assessing deterioration using impairment and functional outcome measures in chronic inflammatory demyelinating polyneuropathy: A post-hoc analysis of the immunoglobulin overtreatment in CIDP trial. *J Peripher Nerv Syst*. 2022;27(2):144-158. doi:10.1111/jns.12497

APPENDIX A: APPENDICES

TABLE A1 Change scores of visits with and without meaningful deterioration according to the secondary patient anchor for the I-RODS, grip strength and MRC-SS (VD − PA and VD + PA, n = 176 visits of 60 patients)

Category of visit according to patient anchor	n visits	I-RODS, centiles (mean change [SD])	I-RODS, MCID-SE (mean change [SD])	Grip strength, kPa (mean change [SD])	MRC-SS, points (median change [IQR])
Visits without meaningful deterioration (VD − PA)	124 visits of 45 patients	1 (SD 7)	0.00 (SD 0.95)	−4 (SD 15)	0.0 (IQR −2 to 0)
Visits with meaningful deterioration (VD + PA)	52 visits of 36 patients	−7 (SD 8)	−1.10 (SD 1.32)	−10 (SD 16)	0.0 (IQR 0–0)

Note: For grip strength, change scores of the hand with the maximum decrease are shown measured using a Martin-Vigorimeter ranging from 0-160 kPa. The MRC-SS used for analyses ranged from 0 (total paralysis) to 60 (normal strength).
Abbreviations: kPa, kilopascal; IQR, interquartile range; I-RODS, Inflammatory Rasch-built Overall Disability Scale; MCID, minimal clinically important differences; MCID-SE, individual change score divided by their corresponding stand error of the difference; MRC-SS, MRC sum score; SD, standard deviation; VD + PA, visit with meaningful deterioration according to the patient anchor; VD − PA, visit without meaningful deterioration according to the patient anchor.

TABLE A2 MCIDs during visits without meaningful deterioration according to the secondary patient anchor (VD − PA, n = 124 visits of 46 patients)

Outcome measure	MCID cutoff for deterioration	No MCID for deterioration (%; n/N visits of n patients)	MCID deterioration (%; n/N visits of n patients)
Singular MCIDs			
I-RODS	−4 centiles	81% (101/124 visits of 38 patients)	19% (23/124 visits of 15 patients)
	−1.96 MCID-SE	100% (124/124 visits of 46 patients)	0% (0/124 visits)
Grip strength	−8 kPa	60% (74/124 visits of 30 patients)	40% (50/124 visits of 25 patients)
	−14 kPa	74% (92/124 visits of 39 patients)	26% (32/124 visits of 20 patients)
MRC-SS	−2 points	94% (116/124 visits of 44 patients)	6% (8/124 visits of 7 patients)
	−4 points	99% (123/124 visits of 45 patients)	1% (1/124 visit)
Combined MCIDs			
I-RODS and either grip strength or MRC-SS	−4 centiles and either −8 kPa or − 2 points	95% (118/124 visits of 43 patients)	5% (6/124 visits of 6 patients)
	−4 centiles and either −14 kPa or − 2 points	96% (119/124 visits of 43 patients)	4% (5/124 of patients of 5 patients)
I-RODS or MRC-SS	−4 centiles or − 2 points	77% (95/124 of 38 patients)	23% (29/124 of 20 patients)
	−1.96 MCID-SE or − 2 points	94% (116/124 of 44 patients)	6% (8/124 of 7 patients)

Note: The table shows the percentages of visits without meaningful deterioration (VD − PA) during which any of the existing MCIDs were reached using change scores between each follow-up visit and the baseline visit. No MCID for deterioration was reached when the change score was >−the MCID value and the MCID for deterioration was reached when the change score was ≤−the MCID value.

Combinations between outcome measures were made by using an "and" and/or "or" function. For the combined MCID of "−4 centile on the I-RODS or −2 points on the MRC-SS", patients had to decrease at least 4 centiles *or* at least 2 points on the MRC-SS to be classified as deteriorated by the combined MCID. For the combination of the "I-RODS and either grip strength or MRC-SS", patients had to decrease at least the MCID on the I-RODS *and* had to decrease at least the MCID on *either* grip strength *or* the MRC-SS.

For grip strength, change scores of the hand with the maximum decrease are shown measured using a Martin-Vigorimeter ranging from 0 to 160 kPa. The MRC-SS used for analyses ranged from 0 (total paralysis) to 60 (normal strength).

Abbreviations: kPa, kilopascal; I-RODS, Inflammatory Rasch-built Overall Disability Scale; MCID, minimal clinically important differences; MCID-SE, individual change score divided by their corresponding stand error of the difference; MRC-SS, MRC sum score; VD + PA, visit with meaningful deterioration according to the patient anchor; VD − PA, visit without meaningful deterioration according to the patient anchor.

TABLE A3 Accuracy of existing MCIDs, combinations of existing MCIDs, and MCIDs optimized for deterioration of the I-RODS, grip strength and MRC-SS

Outcome measure	Treatment anchor AUC (95% CI)	Patient anchor AUC (95% CI)	MCID method	MCID cutoff for deterioration	Treatment anchor Sensitivity (95% CI)	Treatment anchor Specificity (95% CI)	Patient anchor Sensitivity (95% CI)	Patient anchor Specificity (95% CI)
I-RODS, centiles	0.86 (0.77-0.95)	0.90 (0.82-0.97)	Optimized for deterioration based on Youden's index	−3 centiles	92% (79-100)	66% (50-81)	86% (71-96)	79% (64-93)
			Existing MCID	−4 centiles	83% (67-96)	69% (53-83)	79% (64-93)	82% (68-96)
			Optimized for deterioration based on maximized specificity	−8 centiles	63% (42-83)	83% (69-94)	61% (43-79)	93% (82-100)
I-RODS, MCID-SE	0.87 (0.78-0.96)	0.90 (0.83-0.97)	Optimized for deterioration based on Youden's index	−0.515 MCID-SE	96% (79-100)	69% (56-83)	86% (71-96)	82% (68-96)
			Optimized for deterioration based on maximized specificity	−1.82 MCID-SE	63% (42-79)	92% (81-100)	57% (39-75)	100% (100-100)
			Existing MCID	−1.96	50% (29-71)	92% (83-100)	46% (29-64)	100% (100-100)
Grip strength, kPa	0.72 (0.58-0.86)	0.65 (0.50-0.79)	Existing MCID	−8 kPa	73% (55-91)	55% (39-72)	71% (54-89)	57% (39-75)
			Existing MCID	−14 kPa	64% (45-82)	78% (64-92)	46% (29-64)	68 (50-86)
			Optimized for deterioration based on Youden's index	−15 kPa	64% (45-82)	81% (67-92)	46% (29-64)	71% (54-86)
			Optimized for deterioration based on maximized specificity	−15 kPa				
MRC-SS, points	0.69 (0.55-0.83)	0.68 (0.55-0.81)	Optimized for deterioration based on maximized specificity	NA[a]	-	-	-	-
			Existing MCID	−2 points	46% (25-67)	94% (86-100)	36% (18-54)	93% (82-100)
			Optimized for deterioration based on Youden's index	−2 points				
			Existing MCID	−4 points	33% (17-50)	100% (100-100)	25% (11-43)	96% (89-100)
I-RODS and either grip strength or MRC-SS	0.81 (0.70-0.91)	0.71 (0.57-0.83)	Combined MCID	−4 centiles and either −14 kPa or −2 points	70% (47-87)	92% (78-98)	54% (34-72)	89% (72-97)
	0.80 (0.67-0.89)	0.78 (0.66-0.88)	Combined MCID	−4 centiles and either −8 kPa or −2 points	74% (52-90)	83% (67-94)	68% (48-84)	89% (72-98)
I-RODS or MRC-SS			Combined MCID		67% (45-84)	89% (74-97)	57% (37-76)	93% (77-99)

TABLE A3 (Continued)

Outcome measure	Treatment anchor AUC (95% CI)	Patient anchor AUC (95% CI)	MCID method	MCID cutoff for deterioration	Treatment anchor Sensitivity (95% CI)	Treatment anchor Specificity (95% CI)	Patient anchor Sensitivity (95% CI)	Patient anchor Specificity (95% CI)
	0.80 (0.68-0.89)	0.75 (0.62-0.86)		−1.96 MCID-SE or −2 points				
	0.75 (0.62-0.85)	0.80 (0.68-0.90)	Combined MCID	−4 centile or −2 points	88% (68-97)	67% (49-81)	82% (63-94)	79% (59-92)

Note: Calculations are based on data of the last visit of the RCT (n = 60 for the treatment anchor and n = 56 for the patient anchor), using ROC analysis and both anchor as reference standards. The MCID for deterioration was reached when the change score was ≤—the MCID value and no MCID for deterioration was reached when the change score was >—the MCID value.

Combinations between outcome measures were made by using an "and" and/or "or" function. For example, for the MCID of "−4 centile on the I-RODS and −2 points on the MRC-SS", patients had to decrease at least 4 centiles *as well as* at least 2 points on the MRC-SS to be classified as deteriorated by the combined MCID. This means that a patient that deteriorated at least the MCID of −4 centiles on the I-RODS, but did not deteriorate at least the MCID of −2 points on the MRC-SS, was not classified as deteriorated by this combined MCID. For the combined MCID of "-4 centile on the I-RODS or -2 points on the MRC-SS", patients had to decrease at least 4 centiles *or* at least 2 points on the MRC-SS to be classified as deteriorated by the combined MCID. For the combination of the "I-RODS and either grip strength or MRC-SS", patients had to decrease at least the MCID on the I-RODS *and* had to decrease at least the MCID on *either* grip strength *or* the MRC-SS.

The AUC is a measure of the discriminating potential of the outcome measure at various cutoffs. A higher AUC value indicates that the outcome measure is better at distinguishing between patients with and without meaningful deterioration. The AUC of the combined outcome measure is based on the C-statistic. The sensitivity constitutes of the percentage of patients in which meaningful deterioration is correctly identified by the MCID of all patients with meaningful deterioration. The specificity constitutes of the percentage of the patients in which the absence of meaningful deterioration is correctly identified by the MCID of all patients without meaningful deterioration. The confidence intervals of the sensitivity and specificity are based on 2000 stratified bootstrap replicates.

The MCID based on Youden's index was selected based on the cutoff with the highest result of the following formula: sensitivity + specificity −1. For calculation of the MCID based on the maximized specificity, the cutoff with the highest specificity was selected with a minimum sensitivity of 60%.

For grip strength, change scores of the hand with the maximum decrease are shown measured using a Martin-Vigorimeter ranging from 0 to 160 kPa. The MRC-SS used for analyses ranged from 0 (total paralysis) to 60 (normal strength).

Abbreviations: AUC, area under the curve; CI, confidence intervals; IQR, interquartile range; I-RODS, Inflammatory Rasch-built Overall Disability Scale; kPa, kilopascal; MCID, minimal clinically important differences; MCID-SE, individual change score divided by their corresponding stand error of the difference; MRC-SS, MRC sum score; SD, standard deviation.

[a]Not applicable: the lower limit of 60% sensitivity was not associated with deterioration. The cutoff had therefore no clinical meaning and was not shown.

TABLE A4 Accuracy of combined MCIDs of the I-RODS, grip strength and MRC-SS

Combination of outcome measures	MCID cutoffs for deterioration	Sensitivity (95% CI)	Specificity (95% CI)	Youden's index
I-RODS and grip strength	−1.96 MCID-SE and −14 kPa	35% (16-57)	100% (90-100)	0.35
I-RODS and MRC-SS	−1.96 MCID-SE and −2 points	29% (13-51)	97% (85-99)	0.26
I-RODS and MRC-SS	−4 centile and −2 points	42% (22-63)	97% (85-99)	0.39
I-RODS and MRC-SS	−3 centile and −2 points[a]	42% (22-63)	97% (85-99)	0.39
I-RODS and grip strength	−1.96 MCID-SE and −8 kPa	35% (16-57)	94% (81-99)	0.29
I-RODS and grip strength	−4 centile and −14 kPa	55% (32-76)	94% (81-99)	0.49
I-RODS and grip strength	−3 centile and −15 kPa[a]	55% (32-76)	94% (81-99)	0.49
I-RODS and either grip strength or MRC-SS	−3 centile and either −15 kPa or −2 points[a]	70% (47-87)	92% (78-98)	0.62
I-RODS and either grip strength or MRC-SS	−4 centile and either −14 kPa or −2 points	70% (47-87)	92% (78-98)	0.62
I-RODS or MRC-SS	−1.96 MCID-SE or −2 points	67% (45-84)	89% (74-97)	0.56
I-RODS and grip strength	−4 centiles and −8 kPa	59% (36-79)	83% (67-93)	0.42
I-RODS and either grip strength or MRC-SS	−4 centiles and either −8 kPa or −2 points	74% (52-90)	83% (67-94)	0.57
I-RODS or grip strength	−1.96 MCID-SE or −14 kPa	78% (56-84)	69% (52-84)	0.47
I-RODS or MRC-SS	−4 centile or −2 points	88% (68-97)	67% (49-81)	0.55
I-RODS or grip strength	−1.96 MCID-SE or −8 kPa	87% (66-97)	53% (35-70)	0.40
I-RODS or grip strength	−4 centile or −14 kPa	91% (71-99)	53% (35-70)	0.45
I-RODS or grip strength	−3 centile or −15 kPa[a]	100% (86-100)	53% (35-70)	0.53
I-RODS or grip strength	−4 centile or −8 kPa	96% (79-100)	42% (26-59)	0.38

Note: The gray parts of the table indicate that the sensitivity or specificity is above 60% and the Youden's index above 0.50. The darker areas implicate higher (better) values.

Combinations between outcome measures were made by using an "and" and/or "or" function. For example, for the MCID of "−4 centile on the I-RODS and −2 points on the MRC-SS," patients had to decrease at least 4 centiles *as well as* at least 2 points on the MRC-SS to be classified as deteriorated by the combined MCID. This means that a patient that deteriorated at least the MCID of −4 centiles on the I-RODS, but did not deteriorate at least the MCID of −2 points on the MRC-SS, was not classified as deteriorated by this combined MCID. For the combined MCID of "−4 centile on the I-RODS or −2 points on the MRC-SS," patients had to decrease at least 4 centiles *or* at least 2 points on the MRC-SS to be classified as deteriorated by the combined MCID. For the combination of the "I-RODS and either grip strength or MRC-SS," patients had to decrease at least the MCID on the I-RODS *and* had to decrease at least the MCID on *either* grip strength *or* the MRC-SS.

The Youden's index combines the values of the sensitivity and specificity using the following formula: sensitivity + specificity − 1, and ranges from 0 (worst) to 1 (best).

Abbreviations: CI, confidence intervals; I-RODS, Inflammatory Rasch-built Overall Disability Scale; kPa, kilopascal; max., maximized; MCID, minimal clinically important differences; MCID-SE, individual change score of the I-RODS divided by their corresponding stand error of the difference; MRC-SS, MRC sum score.

[a]MCID cutoff optimized for deterioration calculated based on this sample.