Analysis of relapse by inflammatory Rasch-built overall disability scale status in the PATH study of subcutaneous immunoglobulin in chronic inflammatory demyelinating polyneuropathy

Ingemar S. J. Merkies\textsuperscript{1,2} | Ivo N. van Schaik\textsuperscript{3,4} | Vera Bri\textsuperscript{5,6} | Hans-Peter Hartung\textsuperscript{7,8,9} | Richard A. Lewis\textsuperscript{10} | Gen Sobue\textsuperscript{11} | John-Philip Lawo\textsuperscript{12} | Orell Mielke\textsuperscript{12} | David R. Cornblath\textsuperscript{13} | on behalf of the PATH study group

1Department of Neurology, Maastricht University Medical Center, Maastricht, the Netherlands
2Department of Neurology, Curaçao Medical Center, Willemstad, Curaçao
3Department of Neurology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, the Netherlands
4Board Department of the Hospital, Spaarne Gasthuis, Haarlem, the Netherlands
5Department of Medicine (Neurology), University Health Network, University of Toronto, Toronto, Canada
6Institute for Research and Medical Consultations, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
7Department of Neurology, UKD and Center for Neurology and Neuropsychiatry, LVR Klinikum, Medical Faculty, Heinrich Heine University, Düsseldorf, Germany
8Brain and Mind Centre, University of Sydney, Sydney, Australia
9Department of Neurology, Medical University of Vienna, Vienna, Austria
10Department of Neurology, Cedars-Sinai Medical Center, Los Angeles, California, USA
11Department of Neurology, Nagoya University Graduate School of Medicine, Nagoya, Japan
12Clinical Development Department, CSL Behring, Marburg, Germany
13Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Abstract
Clinical trials in chronic inflammatory demyelinating polyneuropathy (CIDP) often assess efficacy using the ordinal Inflammatory Neuropathy Cause and Treatment (INCAT) disability score. Here, data from the PATH study was reanalyzed using change in Inflammatory Rasch-built Overall Disability Scale (I-RODS) to define CIDP relapse instead of INCAT. The PATH study comprised an intravenous immunoglobulin (IVIG) dependency period and an IVIG (IgPro10 [Privigen]) restabilization period; subjects were then randomized to weekly maintenance subcutaneous immunoglobulin (SCIG; IgPro20 [Hizentra]) 0.2 g/kg or 0.4 g/kg or placebo for 24 weeks. CIDP relapse was defined as \( \geq 1 \) point deterioration in adjusted INCAT, with a primary endpoint of relapse or withdrawal rates. This retrospective exploratory analysis redefined relapse using I-RODS via three different cut-off methods: an individual variability method, fixed cut-off of
1 | INTRODUCTION

Chronic inflammatory demyelinating polyneuropathy (CIDP) is an acquired neuropathy and is thought to have an autoimmune-mediated etiology.® CIDP is characterized by symmetrical, proximal, and distal bilateral weakness and somatosensory alterations in arms and legs that usually worsen over 8 weeks or longer.® The 10-point Inflammatory Neuropathy Cause and Treatment (INCAT) disability score has been used to assess efficacy in many trials.® The INCAT score assesses the functionality of the arms and legs by giving a 0-5 score for arms and for legs, with 0 representing no disability and 5 representing no arm function or inability to stand/walk.® There are several issues with the INCAT scoring system including the greater emphasis on leg dysfunction and the fact that the difference in disability between each numerical value is not equally weighted. While used as a linear scale in trials, a 1-point change may have different clinical significance depending on where on the INCAT scale the change occurs. This is highlighted by the use of the adjusted INCAT score in which changes in upper limb function from 0 to 1 (normal to minor symptoms) or 1 to 0 are excluded as they are often considered not to be clinically relevant.¹⁰

Another potential efficacy measure for neuropathies is the Inflammatory Rasch-built Overall Disability Scale (I-RODS), a 24-item patient-reported outcome measure. This questionnaire captures activity and participation restrictions and disease impact on daily tasks.¹¹ It has been suggested that I-RODS, as an interval metric, may be more suitable to assess the range of disability changes in CIDP and may have greater relevance to clinical practice.¹² This has led to discussions as to whether clinical trials should use I-RODS as a primary endpoint instead of INCAT.¹³,¹⁰

The placebo-controlled Polynueopathy And Treatment with Hizentra (PATH) study investigated subcutaneous immunoglobulin (SCIG; IgPro20 [Hizentra, CSL Behring, King of Prussia, PA, USA]) as maintenance therapy for CIDP. The primary endpoint of the PATH study was CIDP relapse or withdrawal, with relapse defined as a ≥1-point increase (deterioration) in adjusted INCAT score.³ Change in I-RODS was assessed as a secondary outcome in PATH. Before I-RODS can be accepted as a suitable primary outcome measure, the clinimetric behavior of I-RODS in clinical trials must be more extensively investigated, and the optimal cut-off threshold of relapse by I-RODS determined. Here, the relapse or withdrawal rate data from PATH are re-analyzed using changes in I-RODS score for those already INCAT-relapsed to define relapse criteria and to assess whether altering the definition of relapse affects the efficacy outcomes from PATH.³

2 | MATERIALS AND METHODS

2.1 | The PATH study

The methods and primary results of the PATH study (NCT01545076) have been reported.³ Briefly, the PATH study investigated two doses of SCIG (IgPro20) vs placebo as maintenance therapy for CIDP. Before randomization to SCIG or placebo, subjects underwent intravenous immunoglobulin (IVIG) withdrawal as part of an IgG dependency test; subjects who showed clinical deterioration were then attempted to be restabilized with IVIG (IgPro10 [Privigen, CSL Behring]) prior to enrollment to the randomized SCIG period of the study. Clinical deterioration in the IgG dependency test was initially defined as ≥1-point increase in adjusted INCAT score; following a study amendment, a ≥4-point decrease in I-RODS or decrease of ≥8 kPa in grip strength were also included as deterioration criteria to assess IgG dependency. CIDP stability was defined as an adjusted INCAT score returning to at least the score recorded at screening and at two or more consecutive visits in the restabilization phase.

Subjects who were restabilized were then randomized to weekly SCIG maintenance therapy (0.2 or 0.4 g/kg body weight [bw]) or placebo for 24 weeks.³ The primary endpoint of the PATH study was the rate of CIDP relapse (defined as ≥1-point increase in adjusted INCAT score) or withdrawal in the SCIG period. Change in I-RODS, grip strength, and Medical Research Council (MRC) sum scores were captured as secondary efficacy endpoints.

2.2 | I-RODS as relapse criterion methodology

This analysis assessed relapse or withdrawal rates using changes in I-RODS to define relapse rather than changes in INCAT. The
original study design used INCAT deterioration to define relapse, at which point subjects were removed from the study. Therefore, the dataset available for the I-RODS analysis has the caveat that subjects could have been removed from the study due to INCAT deterioration before deterioration by I-RODS could be assessed. These patients are classed as "unknown" rather than assuming relapse status in this analysis. Furthermore, this current analysis could only be undertaken in subjects who had raw I-RODS scores available both at baseline and at least one post-baseline visit. These limitations must be taken into consideration when analyzing the results.

In the SCIG period of the PATH study, relapse according to I-RODS was calculated using three methods: (A) a fixed cut-off of ≥4-point deterioration on I-RODS raw score, with a maximal score of 48, (B) a fixed cut-off of ≥8-points deterioration on I-RODS centile score, or (C) individual variability defined by Draak et al.\textsuperscript{14}

For the raw score method, intra-subject differences of the post-baseline values to baseline were derived. Regarding the raw score method, deterioration was classed as ≤−4-point change in I-RODS. For the centile score method, subjects were classified as deteriorated (relapsed) if the change in I-RODS was ≤−8 points. The thresholds were chosen for both the centile score and raw score methods based on MCID calculations, although these are yet to be fully validated for I-RODS.

For the individual variability method, raw I-RODS scores were Rasch-transformed, and intra-subject differences between post-baseline visits and baseline were calculated ($\text{diff}_{ij}$, where $i$ is the subject and $j$ is the visit). For each post-baseline visit, the SE (SE) of the difference was calculated by taking paired values into account only (SE). Minimal clinically important difference (MCID) per subject at each post-baseline visit was then derived as MCID ($\text{MCID}_{ij} = \text{diff}_{ij} / \text{SE}_j$). Subjects were classified as deteriorated (relapsed) if the MCID was ≤−1.96.

### 2.3 Statistical analysis

All statistical analyses were performed using SAS software. These included the following:

- Deriving the 95% Wilson score confidence intervals for the proportions of I-RODS relapse or withdrawal by individual variability method, centile score method and raw score method, and proportions of INCAT relapse or withdrawal rates in the reduced I-RODS analysis population
- Deriving the Kaplan-Meier failure probability estimates, and performing the log-rank test for the $P$ values, for the time to I-RODS relapse or withdrawal by individual variability method, centile score method and raw score method, and time to INCAT relapse or withdrawal rates in the reduced I-RODS analysis population

All confidence intervals and $P$ values are unadjusted.

### 3 RESULTS

A total of 172 subjects demonstrated CIDP stability by adjusted INCAT at the end of the restabilization period and were randomized to placebo (N = 57), 0.2 g/kg IgPro20 (N = 57) or 0.4 g/kg IgPro20 (N = 58).

Raw I-RODS scores were available at baseline and at least one post-baseline visit in 45 subjects (79%) in the placebo group, 50 (88%) of the 0.2 g/kg IgPro20 group, and 52 (91%) of the 0.4 g/kg IgPro20 group and were therefore used for the I-RODS relapse rates re-analysis. While this analysis was not designed as a direct comparison of INCAT and I-RODS, for context, in this smaller
subset, the relapse or withdrawal rate using adjusted INCAT was 60% for placebo, 38% for 0.2 g/kg IgPro20 and 19% for 0.4 g/kg IgPro20.

When relapse or withdrawal rates in the SCIG phase were assessed with relapse defined by a change in I-RODS status, with patients having INCAT relapse prior to I-RODS relapse counted as “unknown”, outcomes were similar using all three methods to define I-RODS relapse, with higher relapse rates seen for placebo compared with the IgPro20 groups (Figure 1). Using the raw score, relapse or withdrawal rates were 47% for placebo, 34% for 0.2 g/kg IgPro20 and 19% for 0.4 g/kg IgPro20. Using the centile score, relapse or withdrawal rates were 40%, 28%, and 15%, respectively for placebo, 0.2 g/kg IgPro20 and 0.4 g/kg IgPro20 and using the individual variability method were 49%, 40%, and 27%, respectively.

A total of 2 patients in the placebo group, 5 patients in the 0.2 g/kg IgPro20 group, and 5 patients in the 0.4 g/kg IgPro20 group completed or withdrew from the study without INCAT relapse, but with relapse according to I-RODS centile score.

In terms of time to relapse, there was no clear trend of one method (INCAT or I-RODS) being able to identify relapse significantly earlier than the others (P values >0.05; Figure 2), therefore it could be likely that those with “unknown” I-RODS relapse status who relapsed based on INCAT would have relapsed on I-RODS within the study timeframe if they remained in the study.

Furthermore, an analysis of time to relapse in 31 patients who relapsed according to both methods found no relevant difference in timing of relapse between the two methods; 71% of relapses occurred on the same day for both methods (Figure 3).
When assessing relapse or withdrawal rates with relapse defined by I-RODS, the results were similar to the original adjusted INCAT relapse or withdrawal rates seen in the PATH study. Our results were similar for all three methods of I-RODS analysis, suggesting that I-RODS data can be used in the raw, but preferably in the centile score format. While the individual variability or centile score methods may be better from a statistical analysis point of view, they are not as readily accessible as ordinal scores, as centiles need to be calculated and the variability method involves several formulae.

A limitation to the current PATH study analysis to consider when interpreting the results is the fact that, as the study originally used adjusted INCAT as the primary endpoint, some subjects were withdrawn from the study following relapse according to adjusted INCAT but before they relapsed according to I-RODS. Further data on adjusted INCAT relapsers following withdrawal was therefore not available so it was not possible to determine whether they subsequently relapsed according to I-RODS.

INCAT is a standard choice for primary endpoints in CIDP trials, and from our original analysis, was shown to be effective at showing a difference in response between placebo and the two doses of IgPro20. While our results show a similar outcome for adjusted INCAT and I-RODS, some physicians believe that I-RODS could also be an appropriate method of efficacy assessment due to its interval nature and increased scope of disability assessment. While I-RODS provides additional functional information compared with INCAT, this study was not primarily designed for comparison purposes, and therefore a comparison of the two disability scores is beyond the scope of this work. When designing any clinical trial, it is important to choose the most appropriate outcome measure in order to define efficacy. INCAT has been criticized as this measure only assesses functional aspects of arms and legs, and it has been postulated that INCAT may underestimate the full scope of impacts that neuropathies have on daily life. However, despite this criticism INCAT has been demonstrated to be responsive in various trials. The INCAT questions regarding arm function have been criticized for assessing similar aspects of activities of daily living, which although differently questionned, partly overlap and embrace the same concept. The limitation of these questions is that they generally all rely on the same functional ability, so issues in one question will often be reflected in the other questions. A change in INCAT points is not as easily observed for the leg scores, which relate to a more discrete deterioration in function. A small change in the arm scores could therefore skew the results for INCAT. In PATH, adjusted INCAT relapses were mainly driven by arm scores (data not shown).

Some physicians have questioned the subjective nature of the I-RODS measure, as it could be argued that a patient’s mood could affect responses to the questionnaire. However, a recent study showed a strong correlation between I-RODS scores and grip strength in patients with neuropathies, supporting the correlation between subjective and objective measures. In addition, in a study by Vanhoutte et al. of patients with CIDP or Guillain-Barré syndrome (GBS), responders were assessed through three outcome measures (I-RODS, Rasch-transformed MRC score or Rasch-transformed modified INCAT sensory score) using the MCID method with a cut-off of 1.96. At the end of the 12-month follow-up period, more patients showed a meaningful change in I-RODS (GBS 97.1%, CIDP 42.1%) than MRC score (GBS 70.6%, CIDP 29%) or INCAT sensory score (GBS 35.3%, CIDP 21.6%). This suggests that I-RODS can be used as a clinically sensitive outcome measure in CIDP and other neuropathies. These findings were validated by the external criterion responsiveness method (capturing patients’ voice), showing a higher correlation for the I-RODS scores, compared with the other metrics. A study by Stucki et al. exploring the generation of change scores in the physical ability scale of the SF-36, concluded that numerically equal gains may differ in their meaning depending on baseline health status, and pointed out challenges associated with the interpretation of change scores in ordinal clinical scales. Limitations and fundamental deficiencies associated with the use of ordinal level scales are discussed in detail in Merbitz et al.

The present study does not allow a comparison of the sensitivity of INCAT and I-RODS in their ability to detect relapse in CIDP. Given that there were more patients who relapsed based on INCAT, compared with I-RODS, one could argue that INCAT is more sensitive. While it’s possible that INCAT is more sensitive than I-RODS, the higher number of patients who relapsed on INCAT could be attributed to methodological issues related to study design, as patients who relapsed by I-RODS would not be detected if I-RODS relapse occurred after INCAT relapse.

In summary, IgPro20 was shown to be efficacious as a maintenance treatment for CIDP when relapse was defined using I-RODS instead of adjusted INCAT, further validating the original INCAT results. All three I-RODS methods used in this analysis were found to be comparable and could be considered instead of, or as a validation of, adjusted INCAT to assess efficacy in CIDP. However, prospective comparison between these two outcome measures or a combination of both in terms of capturing responsiveness could be performed before firm conclusions can be drawn.
I. N. van Schaik chairs a steering committee for CSL Behring and received departmental honoraria for serving on scientific advisory boards for CSL Behring. He received departmental research support from The Netherlands Organization for Scientific Research and from the Dutch Princes Beatrix Fonds. All lecturing and consulting fees for INS were donated to the Stichting Klinische Neurologie, a local foundation that supports research in the field of neurological disorders. He is a member of the Scientific Board of the Kreuth III meeting on the optimal use of plasma-derived medicinal products, especially coagulation factors and normal immunoglobulins organized under the auspices of the European Directorate for the Quality of Medicines & HealthCare (EDQM).

V. Bril is a consultant to CSL Behring, Grifols, Pfizer, UCB, Bionevia, and ArgenX, and has received research support from Baxalta (Shire), CSL Behring, Grifols, Bionevia UCB, and ArgenX.

H-P. Hartung has acted on and data monitoring committees for Bayer Healthcare, Biogen, Celgene BMS, CSL Behring, GeNeuro, Merck, Novartis, Octapharma, Receptos, Roche, Sanofi Genzyme, TG Therapeutics and VielaBio and advisory boards for Alexion and Lundbeck.

R. A. Lewis has received consultation fees and/or served on scientific advisory boards for CSL Behring, Axelacare Health Solutions, Pharmnext, Biotest, Kedrion, NuFactor Inc., Optioncare and Grifols.

G. Sobue has received funds from Mitsubishi Tanabe Pharma Corporation, Takeda Pharmaceutical Company Limited, and CSL Behring.

J-P Lawo and O. Mielke are employees of CSL Behring.

D. R. Cornblath has acted as a consultant for Acetylon Pharmaceuticals Inc., Alnylam Pharmaceuticals, Annexon Biosciences, Akros Pharma, argenx BVBA, Biotest Pharmaceuticals Inc., Boehringer Ingelheim, Cigna Health Management, Inc., CSL Behring, DP Clinical, Inc., Grifols S.A., Hansa Medical Inc., Karos Pharmaceuticals Inc., Merrimack Pharmaceuticals, Inc., Neurocrine Biosciences, Novartis Corp., Octapharma AG, Pharmnext SAS, Seattle Genetics, Inc., Sun Pharmaceuticals and Syntimmune. He has acted on a data safety monitoring board for Pfizer Inc., Ionis Pharmaceuticals, Axovant Sciences LTD., Ampio Pharmaceuticals, PedPharma, Momenta Pharma, and Sanofi. He has a technology license with Acetylon Pharmaceuticals Inc., Akros Pharma, AstraZeneca Pharmaceuticals, LP, Calithera Biosciences, Genentech Inc., Karos Pharma, Neurocrine Biosciences, Merrimack Pharmaceuticals, Inc., Seattle Genetics, Inc. and Shire Development, LLC. He serves on the board of directors for The Peripheral Nerve Society and acts on the medical advisory board for GBsCIDP Foundation International.

AUTHOR CONTRIBUTIONS

The PATH study was designed and conceived by Ivo N. van Schaik, Vera Bril, Hans-Peter Hartung, Richard A. Lewis, Gen Sobue, Orell Mielke, David R. Cornblath and Ingemar S. J. Merkies. John-Philip Lawo wrote the statistical analysis plan. Vera Bril and Ivo N. van Schaik collected data. Ivo N. van Schaik, Vera Bril, Ingemar S. J. Merkies, Orell Mielke and John-Philip Lawo reviewed and initially analyzed data. Ingemar S. J. Merkies and John-Philip Lawo reanalyzed data focusing on I-RODS. All authors interpreted data. Ingemar S. J. Merkies drafted the manuscript which was critically reviewed by all other authors. All authors read and approved the final manuscript before submission.

ETHICAL STANDARDS

The experiments in the original (PATH) study were undertaken with the understanding and written consent of each subject, the study conformed to the World Medical Association Declaration of Helsinki published on the website of the Journal of American Medical Association. The study protocol was approved by the ethics committees of all participating centers.

INDIVIDUAL PATIENT DATA SHARING STATEMENT

CSL will only consider requests to share Individual Patient Data (IPD) that are received from systematic review groups or bona-fide researchers. CSL will not process or act on IPD requests until 12 months after article publication on a public website. An IPD request will not be considered by CSL unless the proposed research question seeks to answer a significant and unknown medical science or patient care question. Applicable country-specific privacy and other laws and regulations will be considered and may prevent sharing of IPD.

Requests for use of the IPD will be reviewed by an internal CSL review committee. If the request is approved, and the researcher agrees to the applicable terms and conditions in a data-sharing agreement, IPD that has been appropriately anonymized will be made available. Supporting documents including study protocol and Statistical Analysis Plan will also be provided.

For information on the process and requirements for submitting a voluntary data-sharing request for IPD, please contact CSL at clinicaltrials@cslbehring.com.

ORCID

Ingemar S. J. Merkies https://orcid.org/0000-0002-8516-2361
Ivo N. van Schaik https://orcid.org/0000-0002-9135-6664
Vera Bril https://orcid.org/0000-0002-5805-4883
Hans-Peter Hartung https://orcid.org/0000-0002-0614-6989
Gen Sobue https://orcid.org/0000-0003-4769-5922
John-Philip Lawo https://orcid.org/0000-0002-4321-176X
Orell Mielke https://orcid.org/0000-0003-4047-8175
David R. Cornblath https://orcid.org/0000-0003-2761-489X

REFERENCES

1. Berger M, McCallus DE, Lin CS. Rapid and reversible responses to IVIG in autoimmune neuromuscular diseases suggest mechanisms of action involving competition with functionally important autoantibodies. J Peripher Nerv Syst. 2013;18(4):275-296.
2. Mathey EK, Park SB, Hughes RA, et al. Chronic inflammatory demyelinating polyradiculoneuropathy: from pathology to phenotype. J Neurol Neurosurg Psychiatry. 2015;86(9):973-985.
3. van Schaik IN, Bril V, van Gelooven 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.
4. Terhoeven P, Seybold J, Utz KS, Nickel FT, Lee DH, Linker RA. Longer-term effects of intravenous immunoglobulin treatment in
chronic inflammatory demyelinating polyneuropathy: who benefits? J Neurol Sci. 2020;419:117169.
5. Nobile-Orazio E, Pujol S, Kasiborski F, et al. An international multicenter efficacy and safety study of iQYmune in initial and maintenance treatment of patients with chronic inflammatory demyelinating polyradiculoneuropathy: PRISM study. J Peripher Nerv Syst. 2020;25(4):356-365.
6. Leger JM, De Bleecker JL, Sommer C, et al. Efficacy and safety of Privigen(R) in patients with chronic inflammatory demyelinating polyneuropathy: results of a prospective, single-arm, open-label phase III study (the PRIMA study). J Peripher Nerv Syst. 2013;18(2):130-140.
7. Hughes RA, Donofrio P, Bril V, et al. Intravenous immune globulin (10% caprylate-chromatography purified) for the treatment of chronic inflammatory demyelinating polyradiculoneuropathy (ICE study): a randomised placebo-controlled trial. Lancet Neurol. 2008;7(2):136-144.
8. Allen J, Gelinas D, Lewis R, Nowak R, Wolfe G. Optimizing the use of outcome measures in chronic inflammatory demyelinating polyneuropathy. US Neurol. 2017;13(1):26-34.
11. 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.
12. Vanhoutte EK, Draak TH, Gorson KC, et al. Impairment measures vs inflammatory RODS in GBS and CIDP: a responsiveness comparison. J Peripher Nerv Syst. 2015;20(3):289-295.
13. Mielke O, Bril V, Comblath 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.
14. Draak TH, Vanhoutte EK, van Nes SI, et al. Changing outcome in inflammatory neuropathies: Rasch-comparative responsiveness. Neurology. 2014;83(23):2124-2132.
15. Draak TH, Gorson KC, Vanhoutte EK, et al. Does ability to walk reflect general functionality in inflammatory neuropathies? J Peripher Nerv Syst. 2016;21(2):74-81.
16. Merkies IS, Schmitz PI, van der Meche FG, Samijn JP, van Doorn PA. Clinimetric evaluation of a new overall disability scale in immune mediated polyneuropathies. J Neurol Neurosurg Psychiatry. 2002;72(5):596-601.