Research Article

For reprint orders, please contact: info@becaris.com



Journal of **Comparative Effectiveness Research**

Matching-adjusted indirect treatment comparisons in chronic inflammatory demyelinating polyneuropathy: interpreting current evidence and enhancing future studies

Batyrkhan Kuatov*[,1], Barkha P Patel[2], Muhaimen Siddiqui[2], Diana Tran[2], Paul Spin[2], Laurence Undreiner[3] & Shicheng Weng[4]

[1]CSL Behring, Bern, Switzerland

[2]EVERSANA®, Victoria, British Columbia, V8W 2E1, Canada

[3]CSL Behring™, Paris, France

[4]CSL Behring, Waltham, MA 02451, USA

*Author for correspondence: Batyrkhan.Kuatov@cslbehring.com

Aim: In the absence of direct evidence, indirect treatment comparisons (ITCs) serve as a tool for comparing treatment efficacy to guide evidence-based healthcare decisions. As an example, given the growing number of treatments available for chronic inflammatory demyelinating polyneuropathy (CIDP), it is important to accurately estimate relative treatment effects to ensure the best patient outcomes. Using CIDP as an example, the aim of this study was to review published best practices and guidelines to determine key methodological considerations for conducting robust ITCs. **Materials & methods:** ITCs in CIDP were identified and evaluated against best practices and guidelines, including the National Institute for Health and Care Excellence Technical Support Document 18, the Health Technology Assessment guidelines for assessment of direct and indirect comparisons, and the Preferred Reporting Items for Systematic Reviews and Meta-analyses extension statement for the reporting of systematic reviews and incorporating network meta-analyses. **Results:** Recently published ITCs in CIDP were limited, with two conference materials identified. Both studies utilized matching-adjusted indirect treatment comparisons to assess the comparative efficacy of CIDP therapies. As per best practices and guidelines reviewed, key methodological considerations associated with comparing CIDP pivotal trials to conduct ITCs were identified. Differences in trial target populations and baseline characteristics, heterogeneity in study designs and outcome definitions, and inappropriate pooling of therapies may limit comparability between trials if not carefully controlled for using available frameworks. **Conclusion:** In the absence of direct comparative evidence, ITCs are necessary and useful tools to compare treatment efficacy. Available frameworks can guide ITC feasibility and limit heterogeneity, to ensure robust and reliable ITC outcomes.

Plain language summary

What is this article about? Indirect treatment comparison (ITC) is a statistical method used to compare different therapies for the same disease and can be used to guide treatment recommendations for patients. The use of ITCs is particularly relevant for rare diseases, such as chronic inflammatory demyelinating polyneuropathy (CIDP), that have multiple treatment options, but no clinical trials testing therapies side by side to determine which option works best. This study examined ITCs in CIDP to demonstrate important methodological requirements to ensure unbiased comparisons.

What methodology/protocol is described? A review of how ITCs were used in CIDP studies was conducted and then compared with accepted best-practice guidelines.

What were the results? Several issues were identified in two ITCs conducted in CIDP, including differences between treatments and their associated clinical trials, as well as inappropriate use of certain statistical methods and lack of transparent reporting. These challenges have the potential to limit comparability between treatments if not carefully controlled by following appropriate guidelines.

Becaris PUBLISHING

Why is this important? This study highlights the importance of using ITCs correctly to ensure accurate comparison of different treatments when direct comparative evidence is limited.

First draft submitted: 29 May 2026; Accepted for publication: 2 September 2026; Published online: 9 October 2026

Keywords: case study • chronic inflammatory demyelinating polyneuropathy • CIDP • indirect treatment comparison • ITC • MAIC • matching-adjusted indirect comparison

In the absence of direct comparative evidence, indirect treatment comparisons (ITCs) represent a useful method for comparing treatments and guiding evidence-based healthcare decision-making [1,2]. They are widely used and increasingly accepted by health technology assessment (HTA) agencies, particularly with rare diseases and orphan drug submissions, where direct evidence is limited [3]. In the EU, the Joint Clinical Assessment permits the use of ITC evidence, with guidelines emphasizing rigorous trial selection and evidence synthesis methods [4].

ITCs can be anchored or unanchored and include network meta-analyses (NMAs) and population-adjusted methods (e.g., matching-adjusted indirect comparisons [MAICs]) [5]. To ensure robust and valid ITCs, three key assumptions must be considered: 1) included trials are comparable in terms of potential treatment effect modifiers (TEMs; e.g., trial or patient characteristics), 2) there is no relevant heterogeneity in outcomes across trials and 3) there is no relevant discrepancy or inconsistency between direct and indirect evidence [1]. Detailed guidelines discussing ITC best practices can help minimize bias and increase validity [6–9].

For rare diseases, such as chronic inflammatory demyelinating polyneuropathy (CIDP), ITCs are particularly relevant when there are multiple existing and emerging therapies, but no direct head-to-head clinical trials. With an estimated global incidence of less than 1 per 100,000 per year, CIDP is a rare autoimmune disease characterized by the demyelination and damage of peripheral nerves [10,11]. Therapies for CIDP include immunoglobulins (IGs), administered intravenously (IVIG) or subcutaneously (SCIG), corticosteroids, plasma exchange and the neonatal Fc receptor antagonist efgartigimod (Vyvgart/Vyvgart Hytrulo) (Supplementary Table A1) [11]. Current CIDP treatment guidelines recommend corticosteroids, plasma exchange, or IVIG for induction therapy and either corticosteroids, IVIG, or SCIG for maintenance therapy to prevent relapse [12]. In recent years, SCIG has been increasingly utilized as maintenance therapy due to ease of administration, more favorable adverse-event (AE) profile and steady effect, versus IVIG [13]. The comparative efficacy of these therapies is not well understood.

The current CIDP treatment landscape highlights the importance of ITCs. The objective of this study was to use CIDP as an example to identify key methodological considerations and requirements for generating robust comparative efficacy assessments.

Materials & methods

Published best practice guidelines consulted for this study included the National Institute for Health and Care Excellence (NICE) Technical Support Document 18 (TSD 18) guideline on population-adjusted ITCs, HTA guidelines for the assessment of direct and indirect comparisons, and the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) extension statement for the reporting of systematic reviews and incorporating NMAs (PRISMA-NMA) [6–9]. While the PRISMA-NMA is typically intended for NMAs, it can be adapted to evaluate the quality of evidence from MAICs (Supplementary Material) [7]. Where feasible, the PRISMA-NMA was applied and complemented by NICE TSD 18 recommendations to determine adherence to best practices in terms of identification and justification of TEMs and population alignment, statistical adjustment methods and diagnostics, and transparency and completeness of reporting.

A targeted search was conducted (September 2025) in PubMed to identify ITCs in CIDP using the following terms: ("Polyradiculoneuropathy, Chronic Inflammatory Demyelinating"[MeSH] OR "chronic inflammatory demyelinating polyneuropathy"[tiab] OR CIDP[tiab]) AND ("indirect treatment comparison"[tiab] OR "indirect comparison"[tiab] OR "network meta-analysis"[tiab] OR "mixed treatment comparison"[tiab] OR NMA[tiab] OR "matching-adjusted indirect comparison"[tiab] OR MAIC[tiab] OR "simulated treatment comparison"[tiab] OR STC[tiab] OR "population-adjusted indirect comparison"[tiab]). The search was intended to identify the most relevant evidence.

An additional search of member-restricted abstract and presentation databases was conducted using the authors' paid society memberships, including the American Academy of Neurology (AAN), Peripheral Nerve Society

(PNS), International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and European Academy of Neurology (EAN). These memberships provided access to archived abstracts and presentations; they did not influence abstract selection or acceptance decisions. Methodology utilized in identified ITCs was then evaluated against best practice guidelines.

Results

Literature search

Two conference materials were identified, both utilizing MAICs to compare efgartigimod with IVIG and SCIG therapies for CIDP. The first study was published as a 2024 ISPOR poster comparing the efficacy of efgartigimod with various IG therapies for CIDP [14]. The second study was presented at both the 2025 PNS annual meeting and the 2025 AAN annual meeting comparing an IVIG (Panzyga) with efgartigimod [15,16]. Pivotal clinical trials included in these MAICs: PATH (SCIG, Hizentra; NCT01545076), ADVANCE-CIDP 1 (Hyaluronidase-facilitated SCIG [fSCIG], HyQvia; NCT02549170), PRIMA (IVIG, Privigen; NCT01184846), ADHERE (efgartigimod; NCT04281472) and ProCID (IVIG, Panzyga; NCT02638207). Supplementary Table A2 contains details regarding each trial and Supplementary Table A3 contains a summary of end point measures used in each trial.

MAIC of efgartigimod versus IGs for CIDP

Methods & results

This MAIC indirectly compared the efficacy of efgartigimod with IVIG and SCIG/fSCIG therapies for CIDP [14]. Following a systematic literature review to identify relevant efficacy data, an ITC feasibility assessment was used to determine if the selected trials were sufficiently comparable with the efgartigimod pivotal trial, ADHERE. Four studies were included in the MAIC: ADHERE, PRIMA, PATH and ADVANCE-CIDP 1 (Supplementary Table A2 for clinical summaries and Supplementary Table A4 for summary of ITC results) [14]. An unanchored MAIC using individual patient data (IPD) from Stage A of ADHERE was used to compare the efficacy of efgartigimod with IVIG as induction therapy. The effective sample size (ESS) was low (2.3–4.7), limiting the reliability of the MAIC [8].

Anchored MAICs comparing efgartigimod with SCIG/fSCIG maintenance therapies were conducted using published data [17,18] and IPD from Stage B of ADHERE. A subset of the ADHERE trial population was used to align with PATH and ADVANCE-CIDP 1 eligibility criteria. The authors noted that ADHERE IPD was matched to the baseline characteristics (TEMs) of each compared cohort; however, these characteristics were not reported. Three sets of MAICs were performed to indirectly compare efgartigimod with SCIG (0.2 and 0.4 g/kg) and fSCIG, except for mean grip strength (MGS; MAICs were conducted vs fSCIG only). The MAIC results were then pooled via meta-analysis. In the analyses of efgartigimod versus SCIG/fSCIG products, following the weighting process, the ESS ranged from 46.6 to 84.0 patients [14].

Key end points assessed across trials included adjusted inflammatory neuropathy cause and treatment (aINCAT), Inflammatory Rasch-built Overall Disability Scale (I-RODS) and MGS score [14]. Pairwise MAIC results comparing efgartigimod versus SCIG or fSCIG were inconsistent across outcomes and SCIG dose. However, pooled meta-analysis results showed significant improvement on most outcomes with efgartigimod versus SCIG/fSCIG therapies. One exception was time to aINCAT deterioration, which did not differ between efgartigimod and SCIG/fSCIG therapies in either dose-specific MAICs or the pooled analyses (hazard ratio [HR]: 0.79: [95% CI: 0.44, 1.45]). Key departures from best-practice guidelines are outlined below.

Areas of methodological misalignment with best practice guidelines

Differences in patient populations & baseline characteristics

Guidelines on the conduct and reporting of MAICs state that for anchored MAICs, all relevant TEMs should be accounted for in the analysis. The status of each variable as a TEM should be justified (via literature review, consultation with experts, or quantitative analysis) and specified *a priori* [8,9]. Published results from the MAIC comparing efgartigimod with IG did not report key TEMs and adjustment was limited to those reported in both ADHERE and the comparator study [14], making it impossible to determine whether all relevant TEMs were accounted for. If all TEMs were not addressed, residual confounding could remain [9,19,20]. Reporting gaps were also identified in the ADHERE participants selected for analysis. Specifically, the authors used a matched subset of ADHERE for each comparison to more closely approximate each trial's eligibility criteria [14]. This is a common step performed prior to estimating MAIC weights used for population-adjustment [21]. However, the authors did not state which eligibility criteria were targeted for matching. Eligibility criteria differed across ADHERE, PATH

and ADVANCE-CIDP 1 and it was not stated how those differences were overcome. PATH patients did not require a minimum INCAT score for entry, indicating a source of non-overlap between populations that cannot be addressed by matching the ADHERE trial to PATH. Furthermore, baseline characteristics indicated that PATH participants were healthier at baseline (e.g., I-RODS and MGS were greater and INCAT was lower in PATH vs ADHERE; Table 1); thus, it is possible that PATH participants would see less improvement versus patients with worse symptoms.

Response to IVIG was a requirement for enrollment in PATH and ADVANCE-CIDP 1, whereas ADHERE permitted treatment-naive patients and patients with prior exposure to corticosteroids or IG [17,18,22]. Although there was a subset of IG-dependent patients in ADHERE, tests of IG dependency were conducted differently between trials. IG-dependent patients in ADHERE were identified by deterioration upon IG withdrawal, while IG-dependent patients in PATH were identified by both deterioration upon IG withdrawal and re-stabilization upon IG reinstitution. IG dependency was not tested in ADVANCE-CIDP 1. All three trials restricted maintenance phases to responders, despite differences in IG dependency definitions and eligibility criteria.

There were also differences in baseline characteristics between trials. ADHERE participants had greater baseline disability than those in PATH and ADVANCE-CIDP 1, as indicated by higher baseline INCAT, lower baseline MGS, and lower baseline I-RODS (Table 1). Notably, mean baseline I-RODS scores were ≥7.4 points lower in ADHERE versus PATH and ADVANCE-CIDP 1, exceeding the minimal clinically important difference as defined by an minimal clinically important difference-standard error (SE) ≥1.96 [23]. Poor baseline function in ADHERE patients may lead to greater relative efficacy by providing participants with less opportunity for deterioration; this may result in more favorable results versus trials with greater baseline function, especially when change from baseline (CFB) is used [24]. In addition to baseline disability, there were differences in prior IG-experience between trials (Table 1). No post-adjustment balance statistics (e.g., standardized differences) were reported to enable assessment of the impact of adjustment on between-study population differences.

Table 1. Differences in baseline characteristics across pivotal trials for efgartigimod, SCIG, and fSCIG.

Baseline characteristics[†]	ADHERE Stage B (N = 111)	SCIG/fSCIG comparison; n (%)		
		PATH; IgPro 20 0.2 g/kg (N = 57)	PATH; IgPro 20 0.4 g/kg (N = 58)	ADVANCE – CIDP 1; (N = 62)
Age	54.5 (13.18)	57.5 (12.0)	56.6 (13.6)	55.0 (14.3)
Male, n (%)	73 (66)	42 (74)	31 (53)	36 (58.1)
Race, n (White, %)	73 (66)	NR	NR	58 (93.5)
Typical CIDP, n (%)	97 (87)	NR	NR	NR
Atypical CIDP	14 (13)	NR	NR	NR
Asymmetric	6 (5)	NR	NR	NR
Distal	7 (6)	NR	NR	NR
Pure motor	1 (1)	NR	NR	NR
INCAT score	3.1 (1.51)	2.0 (1.0, 3.0)	2.0 (1.0, 3.0)	3.0 (2.0, 4.0)
I-RODS score	53.6 (17.91)	63.0 (51.0, 73.0)	69.0 (54.0, 80.0)	61.0 (47.0, 73.0)
Grip strength (dominant hand), kPa	54.9 (23.64)	67.0 (56.7, 86.2)	68.4 (46.0, 93.3)	54.0 (42.0, 70.0)
Grip strength (non-dominant hand), kPa	55.4 (28.29)	NR	NR	NR
CIDP treatment within the past 6 months, n (%)				
Corticosteroids	24 (22)	NR	NR	7 (11.3)
Immunoglobulins (IV or SC)	48 (43)	57 (100)	58 (100)	NR
Treatment-naïve	39 (35)	0 (0)[‡]	0 (0)[‡]	NR
IVIG dose during 3 months before screening (g/kg)	NR	2.3 (1.3, 3.0)	2.7 (1.3, 3.4)	NR

[†]Data are mean (SD), median (IQR) or n (%).
[‡]Values are assumed based on all patients having to be IG-dependent in the PATH trial.
CIDP: Chronic inflammatory demyelinating polyneuropathy; fSCIG: Facilitated subcutaneous immunoglobulin; I-RODS: Inflammatory Rasch-Built Overall Disability Scale; IG: Immunoglobulin; INCAT: Inflammatory Neuropathy Cause and Treatment; IQR: Interquartile range; IVIG: Intravenous immunoglobulin; kPa: Kilopascal; MAIC: Matching-adjusted indirect comparison; MCID: Minimal clinically important difference; NR: Not reported; SCIG: Subcutaneous immunoglobulin; SD: Standard deviation.

Inappropriate pooling of SCIG therapies & doses via meta-analysis

The primary conclusions drawn in the analysis comparing efgartigimod with IG were based on the meta-analysis [14]. Pooling provides a single summary result estimate from multiple studies. However, the studies considered for pooling must be sufficiently similar in study designs and populations to accommodate pooling [25]. In this case, comparative results were pooled for distinct SCIG/fSCIG therapies, which may be inappropriate given the heterogeneity across trials (Supplementary Table A2) [14]. For instance, all PATH patients were IG-dependent, while IG dependency was not formally assessed in ADVANCE-CIDP 1 [17,18]. Because ADVANCE-CIDP 1 included patients who may not deteriorate after stopping IG, fewer participants in the placebo arm relapsed. This resulted in a lower placebo relapse rate (31%) versus the IG-dependent PATH population (56%), indicating that the individual MAICs were likely not suitable for pooling due to underlying population differences. In addition to pooling the SCIG/fSCIG therapies, both low- and high-dose SCIG were pooled despite dose-dependent effects reported in PATH [18]. As the higher dose demonstrated greater numerical efficacy, pooling across doses could obscure clinically meaningful differences and limit interpretability.

The pooling of MAIC results also potentially involved a 'double counting' of the ADHERE participants included in each of the individual MAIC analyses, leading to overly precise treatment effects [26]. The authors did not address this possibility, or report attempts to mitigate the impact of double counting patients. Notably, the pairwise MAIC for CFB in aINCAT did not demonstrate a statistically significant difference between efgartigimod and SCIG 0.4 g/kg (MD: -0.58; 95% CI: -1.45, 0.29), whereas the pooled analysis produced a larger point estimate and narrower CI that was statistically significant (MD: -0.85; 95% CI: -1.36, -0.35) [14]. The change in magnitude and precision of the pooled estimate indicates the potential influence of pooling [27]; however, the extent to which this was attributable specifically to double counting cannot be determined from the published data.

MAIC of IVIG (Panzyga) versus efgartigimod for CIDP

Methods & results

This study conducted an unanchored MAIC to indirectly compare the efficacy of IVIG (Panzyga; ProCID) versus efgartigimod (ADHERE) [15,16]. IPD from ProCID were used to compare two doses of IVIG against aggregate data from Stage A (open-label efgartigimod) and Stage B (double-blind, placebo-controlled) of ADHERE [15,16]. For comparisons of IVIG versus open-label efgartigimod (Stage A), IPD from ProCID was re-weighted to match ADHERE Stage A baseline age and sex. Results of this MAIC demonstrated that a higher number of patients treated with either dose of IVIG (1.0 g/kg: 84.7%; 2.0 g/kg: 92.3%) reached confirmed evidence of clinical improvement (ECI) versus those treated with efgartigimod (66.5%) [15,16]. The time to initial confirmed ECI was also significantly shorter for both 1.0 g/kg (3.71 vs 6.14 weeks, HR: 1.85 [95% CI: 1.37, 2.50]) and 2.0 g/kg (3.14 vs 6.14 weeks, HR: 2.39 [95% CI: 1.64, 3.48]) IVIG doses versus efgartigimod [15,16]. Greater improvements in aINCAT score and Medical Research Council (MRC) Sum Score were observed for both IVIG doses versus efgartigimod. There was no difference between treatments in CFB I-RODs score and MGS (non-dominant and dominant) [15,16]. For comparisons of IVIG versus maintenance efgartigimod (Stage B), IPD from ProCID was re-weighted a second time to match ADHERE Stage B baseline age and sex. Results of this MAIC demonstrated that patients who responded to Panzyga maintained their response and did not relapse, whereas 25% of patients who responded to efgartigimod during the open-label period had relapsed by Week 20 of the maintenance period [15,16]. See Supplementary Table A5 for a summary of the MAIC results.

Areas of methodological misalignment with best practice guidelines

The MAIC comparing efgartigimod with IVIG was compared against published best practices. While matching was performed for both Stage A and B of the ADHERE trial, patients appeared to only be matched on age and sex at baseline, despite differences in prior treatment exposure and adjusted INCAT score (Table 2) [15,16]. This imbalance is particularly relevant given that ProCID participants were not treatment-naive (i.e., all were previously treated with corticosteroids or IG), while the ADHERE population was more heterogeneous, including treatment-naive, off-treatment, and previously treated patients. These differences limit population comparability and challenge the validity of cross-trial adjustment. As this was an unanchored MAIC, where both TEMs and prognostic factors (PFs) must be adjusted for, residual confounding may have persisted despite attempts to match baseline characteristics [8].

Additionally, no ESS was reported, which is a key diagnostic required under NICE TSD 18 for population-adjusted ITCs [8]. An ESS indicates how much information is retained after re-weighting and whether the weighted population remains representative of the original sample. Without this metric, it is not possible to assess the

Table 2. Differences in baseline characteristics across ADHERE (efgartigimod) and ProCID (IVIG).

Baseline characteristics[†]	ADHERE Stage A (N = 322)	ADHERE Stage B (N = 111)	ProCID		
			0.5 g/kg Panzyga (N = 34)	1.0 g/kg Panzyga (N = 69)	2.0 g/kg Panzyga (N = 36)
Age, years	54.0 (13.92)	54.5 (13.18)	57 (40–64)	59 (51–67)	61 (49–66)
Male, n (%)	208 (65)	73 (66)	21 (62)	38 (55)	23 (64)
Race, n (White, %)	211 (66)	73 (66)	NR	NR	NR
Typical CIDP, n (%)	268 (83)	97 (87)	33 (97)	62 (90)	32 (89)
Atypical CIDP	54 (17)	14 (13)	1 (3)	7 (10)	4 (11)
Asymmetric	29 (9)	6 (5)	NR	NR	NR
Distal	20 (6)	7 (6)	NR	NR	NR
Pure motor	5 (2)	1 (1)	NR	NR	NR
INCAT score	4.6 (1.67)	3.1 (1.51)	4 (4–5)[‡]	4 (4–5)[‡]	4 (4–5)[‡]
I-RODS score	40.1 (14.67)	53.6 (17.91)	27 (20–32)[§]	25 (21–31)[§]	29 (21–32)[§]
Grip strength (dominant hand), kPa	38.5 (24.18)	54.9 (23.64)	51 (32–68)[¶]	53 (39–78)[¶]	52 (41–64)[¶]
Grip strength (non-dominant hand), kPa	39.0 (24.71)	55.4 (28.29)	52 (28–70)[¶]	54 (38–76)[¶]	54 (38–67)[¶]
CIDP treatment prior to enrollment, n (%)					
Corticosteroids	63 (20)	24 (22)	29 (85)	60 (87)	32 (89)
Immunoglobulins (IV or SC)	165 (51)	48 (43)	5 (15)	9 (13)	4 (11)
Treatment-naïve	94 (29)	39 (35)	0	0	0
IVIG at start of dose elevation, median (range)	NR	NR	0.9 (0.3–1.4)	0.4 (0.3–0.8)	0.4 (0.2–0.7)

[†]Data are mean (SD), median (IQR) or n (%), unless otherwise specified.
[‡]Data are aINCAT, range from 0–10.
[§]Data are I-RODS, range from 0–48.
[¶]Data are max grip strength (kPa), range from 0–160.
CIDP: Chronic inflammatory demyelinating polyneuropathy; I-RODS: Inflammatory Rasch-Built Overall Disability Scale; IG: Immunoglobulin; INCAT: Inflammatory Neuropathy Cause and Treatment; IQR: Interquartile range; IVIG: Intravenous immunoglobulin; kPa: Kilopascal; MAIC: Matching-adjusted indirect comparison; MCID: Minimal clinically important difference; NR: Not reported; SCIG: Subcutaneous immunoglobulin; SD: Standard deviation.

precision of the reported estimates, as recommended by NICE [8]. The omission of ESS and balance statistics therefore limits confidence in the validity and reproducibility of the results.

Finally, the MAIC comparing efgartigimod with IVIG consisted of two ITCs: 1) a comparison of the ProCID dose-evaluation phase (which included a loading dose and maintenance doses) with the induction phase (Stage A) of ADHERE and 2) the ProCID dose-evaluation phase with the maintenance phase (Stage B) of ADHERE [15,16]. The first of these ITCs compared outcomes over different observation periods (≤24 weeks for IVIG vs ≤12 weeks induction of efgartigimod for ADHERE), making it difficult to draw conclusions regarding comparative efficacy. The second ITC compared different populations: participants of the ProCID dose-evaluation phase (regardless of initial response) with efgartigimod responders who entered the maintenance phase. There is evidence that patients transitioning from induction to the maintenance phase of ADHERE demonstrated improved quality of life consistent with induction response, as indicated by improvements in baseline functional measures (e.g., INCAT score, I-RODS score, MGS, etc.). Hence, the patient population in Stage B of ADHERE is not likely to be comparable with ProCID as they differ based on prior response to treatment and disease trajectory. These differences may violate the principle of exchangeability required for valid inference, which requires that the included studies evaluate comparable patient populations and outcomes [7,8].

Discussion

MAICs use IPD from one trial to align baseline characteristics with published aggregate data from another to account for cross-trial heterogeneity and estimate treatment differences [9,21,28,29]. This methodology has been increasingly utilized in HTA submissions, including NICE and Canada's Drug Agency-L'Agence des médicaments du Canada (CDA-AMC), which provide recommendations on conducting MAICs [6,30]. Given the weight that MAICs can have in decision-making, it is critical that they are methodologically rigorous, transparently reported and align with best practice guidelines.

The present study reviewed available ITCs comparing CIDP treatments. Two studies were identified that used MAICs to compare the efficacy of various IGs with efgartigimod for first-line treatment of CIDP [14–16]. After evaluating the methodology and conclusions against best practice guidelines [7–9], several methodological challenges were identified, including inappropriate pooling of low- and high-dose SCIG, inappropriate use of meta-analysis and limited adjustment for sources of heterogeneity across CIDP trials. There were notable between-trial differences that limited comparability. Comparisons of both MAICs against best practice guidelines are summarized in Supplementary Table A6.

Methodological challenges associated with heterogeneity across trials included in both MAICs suggest that inter-trial differences may complicate the feasibility and validity of MAICs in CIDP. As observed in the MAIC comparing efgartigimod with IG, the PATH, ADHERE and ADVANCE-CIDP 1 trials all evaluated maintenance therapy to prevent CIDP relapse; however, patient populations differed [14]. PATH and ADVANCE-CIDP 1 enrolled IVIG-responsive patients who were clinically stable prior to randomization and had milder disability. Conversely, ADHERE enrolled a more heterogeneous maintenance population, including patients who demonstrated treatment dependence (including IG-dependence) during run-in withdrawal as well as treatment-naive and previously treated patients, encompassing a broader range of disease severity. These population differences may be clinically meaningful, as variability in prior treatment history and disease severity may affect relapse risk and response patterns during maintenance, complicating cross-trial comparisons.

As observed in the MAIC comparing efgartigimod with IVIG, differences in baseline characteristics in ADHERE and ProCID, including the type of CIDP, patient age, baseline INCAT score and MGS, also suggest fundamental differences in the patient population between trials [15,16]. In ProCID, patients relapsed during IVIG or corticosteroid wash-out to confirm active and treatment-dependent CIDP, whereas in ADHERE, patients deteriorated during treatment withdrawal (run-in) or had documented evidence of clinically meaningful deterioration before entering the open-label induction phase (Stage A). These factors may further complicate the interpretation of outcomes from indirect comparisons.

Differences in trial design, including follow-up duration and last dose post-observation (LDPO) timing, also introduce variability in MAIC results that may be unrelated to treatment effects. In the MAIC comparing efgartigimod with IG for CIDP, follow-up for relapse or time-to-relapse differed: up to 24 weeks in PATH, up to 32 weeks in ADVANCE-CIDP 1, and up to 48 weeks in ADHERE (Supplementary Table A3) [14]. Continuous outcomes were assessed at each study's final LDPO window: at the LDPO up to Week 24 in PATH, end of treatment (~Week 26) in ADVANCE-CIDP 1, and up to Week 48 in ADHERE. These LDPO windows were not aligned and could influence how CFB outcomes were calculated and which participants contributed data at the final timepoint. Additional variability arises from relapse-related dropout, as continuous outcomes were analyzed only among participants who remained relapse-free. This could introduce collider bias, as both treatment and relapse affect whether a patient continues to provide evaluable data [31]; however, the presence of collider bias cannot be confirmed in these analyses. Together, these differences could bias mean-change and relapse-free estimates independent of true treatment effect.

Differences in study design and outcome definitions further complicate comparisons between the two trials included in the MAIC comparing efgartigimod with IVIG for CIDP [15,16]. ProCID evaluated patients immediately after relapse and measured ECI and sustained response over 24 weeks following IVIG re-induction and maintenance therapy. Conversely, ADHERE evaluated time-to-relapse over ≤48 weeks among patients who had improved on open-label efgartigimod and were then randomized to continue or withdraw treatment. As a result, the ProCID outcomes capture recovery after relapse, whereas the ADHERE results reflect durability of response after initial improvement. The longer follow-up in ADHERE increased the window in which relapse events could occur relative to the 24-week treatment phase in ProCID. In ProCID, patients underwent IVIG re-induction after washout and patients in Stage A of ADHERE had open-label efgartigimod induction. Although differences exist, both these phases were broadly comparable as both studies incorporated pre-randomization or induction phases in which patients first deteriorated and were then re-treated; both also assessed ECI following re-treatment after deterioration.

The MAIC comparing efgartigimod with IG for CIDP was recently evaluated in the CDA-AMC reimbursement recommendation for efgartigimod; the appraisal aligned with the current review [32]. The CDA-AMC noted that while adjustment for confounders and TEMs included all variables reported in both ADHERE and the comparator studies, residual bias due to potential unknown confounders likely remained [33]. The key TEMs and PFs adjusted for were reported to have been informed by clinical expert opinion, which may result in incomplete or biased

adjustments relative to an evidence-based approach. The CDA-AMC noted that while the use of meta-analyses to compare efgartigimod with IG was reasonable given the availability of IG treatment data, this approach assumes homogeneity across studies and does not consider clinical and methodological heterogeneity across trials [33]. These findings are consistent with those reported in the present study and highlight the importance of methodological rigor and reporting to ensure valid ITC results, especially when informing HTA decisions.

Future directions

The present study highlights opportunities to improve the rigor and reporting of future ITCs in CIDP. First, investigators should adhere to established methodological and reporting guidelines [7,8], including transparent reporting of covariate selection, balance diagnostics, weighting, ESS, assumptions and sensitivity analyses. Supplementary Table A6 provides a summary of best practices that can be used for designing and reporting future ITCs. Second, given the heterogeneity of CIDP trials, a formal ITC feasibility assessment should be conducted to identify key design and population differences between trials included in the evidence base and develop proactive strategies to mitigate potential resulting biases (e.g., leveraging IPD to compare outcomes over similar observation periods). Where population adjustment is required, clinically and empirically supported TEMs and PFs should be identified and prioritized *a priori*, and the impact of residual differences that cannot be adjusted for should be explicitly considered. Finally, outcome definitions and analytic approaches should be harmonized wherever possible, particularly for commonly assessed CIDP outcomes, and pooling of estimates derived from overlapping patient populations should be avoided or appropriately accounted for. Collectively, these steps can help identify whether ITCs in CIDP are sufficiently robust to inform decision-making or when substantial heterogeneity or limited reporting warrants more cautious interpretation.

In addition to MAICs, other ITC methods could provide information regarding treatment efficacy in CIDP. An NMA can synthesize broader evidence networks and provide comprehensive comparisons of CIDP therapies, which may offer advantages over MAICs when sufficient data exist [34]. However, the appropriateness of a specific ITC technique is context-dependent and determined by factors such as data quality, evidence availability and the relative balance of benefit and uncertainty associated with each method. A feasibility assessment should inform selection of the most appropriate approach based on available evidence, comparability of data sources and potential sources of heterogeneity or bias [2].

Although the methodological limitations identified in the reviewed MAICs limit confidence in their findings, these challenges do not necessarily preclude the use of ITCs in CIDP. Rather, they highlight the need for careful feasibility assessment and appropriate study selection before conducting indirect comparisons. ITCs may be feasible in situations where trials evaluate comparable patient populations, treatment settings, outcome definitions, and follow-up periods, and where important TEMs and PFs can be adequately aligned or adjusted for. Conversely, when substantial differences in patient populations, treatment settings, outcome definitions, or follow-up periods remain after adjustment, the assumptions required for valid indirect comparisons may not be met and findings should be considered exploratory and interpreted with caution.

Limitations

A key limitation of the present study was the small number of available ITCs in CIDP, with identified analyses reported in conference materials that lacked complete methodological details. The targeted search strategy may have limited the number of identified studies, compared with a more comprehensive systematic literature review. The small number of identified MAICs also precluded formal sensitivity or heterogeneity analyses and may reflect potential publication bias toward positive or feasible studies. Incomplete reporting of covariate selection or weighting constrained the ability to assess residual confounding. However, this study sought to identify challenges related to the validity of existing ITCs, not to empirically assess the magnitude of these potential biases. Despite these limitations, this is the first evaluation of MAICs in CIDP, to our knowledge, conducted using internationally recognized methodological standards.

Conclusion

Recurring sources of heterogeneity across CIDP trials and methodological challenges in the conduct of ITCs limit their interpretation. This study provides a practical framework for conducting robust ITCs in this disease space. In the absence of direct head-to-head data, rigorous methodology and transparent reporting are essential to generate reliable evidence.

Summary points

- In the absence of direct comparative evidence, indirect treatment comparisons (ITCs) represent a useful method for comparing treatments and guiding evidence-based healthcare decision-making.
- The use of ITCs is particularly relevant for rare diseases, such as chronic inflammatory demyelinating polyneuropathy (CIDP), for which there are multiple existing and emerging therapies, but no direct head-to-head clinical trials.
- The objective of this study was to use CIDP as an example to identify key methodological considerations and requirements for generating robust comparative efficacy assessments.
- Following a literature search, two conference materials were identified, both utilizing matching-adjusted indirect treatment comparison methodology to compare efgartigimod with intravenous immunoglobulin and subcutaneous immunoglobulin therapies for the treatment of CIDP.
- Challenges with transparent methodological reporting, controls to account for differences in trial target populations and baseline characteristics, heterogeneity in study designs and outcome definitions, and inappropriate pooling of therapies were identified.
- The findings in the present study demonstrate limitations in previously published ITCs in CIDP and highlight key areas for improving rigor and reporting in ITCs moving forward.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at https://becarispublishing.com/doi/epdf/10.57264/cer-2026-0077

Author contributions

B Kuatov, L Undreiner and S Weng contributed to the conception and design of the work and provided critical review and intellectual input throughout the development of the manuscript. M Siddiqui, D Tran, P Spin and BP Patel contributed to the conception and design of the work, drafting the manuscript, conducting the critical appraisal of the evidence, interpreting the findings, and critically revising the manuscript for important intellectual content. All authors contributed to the overall interpretation of the work, approved the final version of the manuscript, and agree to be accountable for all aspects of the work in accordance with JCER criteria.

Acknowledgments

Alphonse Hubsch contributed to the conception and design of the work and provided intellectual input throughout the development of the manuscript. Editorial support was provided by Georgia Greaves of Bioscript Group, Macclesfield, UK.

Financial disclosure

CSL Behring funded this study and also participated in the development, review, and approval of the publication.

Competing interests disclosure

B Kuatov, L Undreiner and S Weng are employees of CSL Behring. M Siddiqui, D Tran, P Spin and BP Patel are employees of EVERSANA, which received funding from CSL Behring for this study. The authors have no other competing interests or relevant affiliations with any organization or entity with the subject matter or materials discussed in the manuscript apart from those disclosed.

Writing disclosure

Medical writing support was provided by Nathan Gock, MSc, and Cassidy Richardson, PhD, of EVERSANA, Victoria, British Columbia, Canada. Writing support funding was provided by CSL Behring.

Open access

This work is licensed under the Creative Commons Attribution 4.0 License. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

References

Papers of special note have been highlighted as: ● of interest; ●● of considerable interest

1. Kiefer C, Sturtz S, Bender R. Indirect comparisons and network meta-analyses. *Dtsch Arztebl. Int.* 112(47), 803–808 (2015).

2. Tanaka S, Igarashi A, De Moor R *et al.* A targeted review of worldwide indirect treatment comparison guidelines and best practices. *Value Health* 27(9), 1179–1190 (2024).
3. Igarashi A, Tanaka S, De Moor R *et al.* Indirect treatment comparisons in healthcare decision making: a targeted review of regulatory approval, reimbursement, and pricing recommendations globally for oncology drugs in 2021–2023. *Adv. Ther.* 42(1), 52–69 (2025).
4. HTA Coordination Group. Practical guideline for quantitative evidence synthesis: direct and indirect comparisons. (2024). https://health.ec.europa.eu/document/download/1f6b8a70-5ce0-404e-9066-120dc9a8df75_en?filename=hta_practical-guideline_direct-and-indirect-comparisons_en.pdf
5. Macabeo B, Quenéchdu A, Aballéa S, François C, Boyer L, Laramée P. Methods for indirect treatment comparison: results from a systematic literature review. *J. Mark. Access Health Policy* 12(2), 58–80 (2024).
6. Health Technology Assessment Coordination Group. Methodological guideline for quantitative evidence synthesis: Direct and indirect comparisons. HTA CG: European Commission (2024). https://health.ec.europa.eu/document/download/4ec8288e-6d15-49c5-a490-d8ad7748578f_en

• **An health technology assessment guideline for the assessment of direct and indirect comparisons.**

7. Hutton B, Salanti G, Caldwell DM *et al.* The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann. Intern. Med.* 162(11), 777–784 (2015).

• **A guideline to improve the completeness of reporting and incorporation of systematic reviews, meta-analyses and network meta-analyses.**

8. Phillippo DM, Ades AE, Dias S *et al.* NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. Unit NDS: Decision Support Unit, ScHARR, University of Sheffield, Sheffield, England, (2016). http://www.nicedsu.org.uk

• **Technical support document outlining methodology for conducting population-adjusted indirect treatment comparisons (ITCs) for submissions to National Institute for Health and Care Excellence.**

9. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for population-adjusted indirect comparisons in health technology appraisal. *Med. Decis. Making* 38(2), 200–211 (2018).
10. Mathey EK, Park SB, Hughes RAC *et al.* Chronic inflammatory demyelinating polyradiculoneuropathy: from pathology to phenotype. *J. Neurol. Neurosurg. Psych.* 86(9), 973–985 (2015).
11. Rajabally YA. Chronic inflammatory demyelinating polyradiculoneuropathy: current therapeutic approaches and future outlooks. *Immunotargets Ther.* 13, 99–110 (2024).
12. 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. *Eur. J. Neurol.* 28(11), 3556–3583 (2021).
13. Cocito D, Peci E, Torrieri MC, Clerico M. Subcutaneous immunoglobulin in chronic inflammatory demyelinating polyneuropathy: a historical perspective. *J. Clin. Med.* 12(22), 6961 (2023).
14. Celico L, Alves A, Iannazzo S, Arvin-Berod C, Stettner M, Skripuletz T. Indirect treatment comparison of efgartigimod vs immunoglobulins in chronic inflammatory demyelinating polyneuropathy (CIDP). *Value Health.* 27(12 Suppl.), S34 (2024).

•• **This matching-adjusted indirect comparison(MAIC) indicated that efgartigimod had comparable or greater efficacy than immunoglobulins (IGs) for the treatment of chronic inflammatory demyelinating polyneuropathy (CIDP), despite study limitations and limited comparability of studies.**

15. Cornblath DR, Wilson J, Baquie M, Wissmann C, Clodi E. *Intravenous immunoglobulin versus efgartigimod for cidp: Matching-adjusted indirect comparison. P455*. Peripheral Nervous Society (PNS) Annual Meeting, Edinburgh, Scotland (17–20 May 2025).

•• **This MAIC indicated comparable or greater efficacy of IVIGs for the treatment of CIPD compared with efgartigimod.**

16. Cornblath DR. *Broadening horizons in autoimmune diseases: Immunoglobulin therapy in an evolving landscape*. American Academy of Neurology, CA, USA (2025). https://www.aan.com/msa/Public/Events/Details/18309

•• **This MAIC indicated comparable or greater efficacy of IVIGs for the treatment of CIPD compared with efgartigimod.**

17. Bril V, Hadden RDM, Brannagan TH 3rd *et al.* Hyaluronidase-facilitated subcutaneous immunoglobulin 10% as maintenance therapy for chronic inflammatory demyelinating polyradiculoneuropathy: the advance-CIDP 1 randomized controlled trial. *J. Peripher. Nerv. Syst.* 28(3), 436–449 (2023).
18. 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.* 17(1), 35–46 (2018).
19. Phillippo DM, Dias S, Elsada A, Ades AE, Welton NJ. Population adjustment methods for indirect comparisons: a review of National Institute for Health and Care Excellence technology appraisals. *Int. J. Technol. Assess. Health Care* 35(3), 221–228 (2019).
20. Petto H, Kadziola Z, Brnabic A, Saure D, Belger M. Alternative weighting approaches for anchored matching-adjusted indirect comparisons via a common comparator. *Value Health* 22(1), 85–91 (2019).
21. Signorovitch JE, Sikirica V, Erder MH *et al.* Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value Health* 15(6), 940–947 (2012).

22. Allen JA, Lin J, Basta I *et al.* Safety, tolerability, and efficacy of subcutaneous efgartigimod in patients with chronic inflammatory demyelinating polyradiculoneuropathy (ADHERE): a multicentre, randomised-withdrawal, double-blind, placebo-controlled, Phase II trial. *Lancet Neurol.* 23(10), 1013–1024 (2024).
23. Draak TH, Vanhoutte EK, van Nes SI *et al.* Changing outcome in inflammatory neuropathies: rasch-comparative responsiveness. *Neurology* 83(23), 2124–2132 (2014).
24. Vickers AJ, Altman DG. Statistics notes: analysing controlled trials with baseline and follow up measurements. *BMJ* 323(7321), 1123–1124 (2001).
25. Saure D, Schacht A, Kadziola Z, Brnabic AJM. Combination of several matching adjusted indirect comparisons (MAICs) with an application in psoriasis. *Pharm. Stat.* 19(5), 532–540 (2020).
26. Senn SJ. Overstating the evidence: double counting in meta-analysis and related problems. *BMC Med. Res. Methodol.* 9, 10 (2009).
27. Hussein H, Nevill CR, Meffen A *et al.* Double-counting of populations in evidence synthesis in public health: a call for awareness and future methodological development. *BMC Public Health* 22(1), 1827 (2022).
28. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med. Decis. Making* 33(5), 607–617 (2013).
29. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J. Clin. Epidemiol.* 50(6), 683–691 (1997).
30. Canada's Drug Agency. Methods guide for health technology assessment (2025). https://www.cda-amc.ca/sites/default/files/MG%20Methods/MG0030-Quantitative-Methods-Manual_Mar%202025.pdf
31. Levy NS, Kezios KL. The same but different?: a systematic review of the impact of selection and collider bias on internal validity. *Epidemiology* 36(4), 473–481 (2025).
32. Canada's Drug Agency. Reimbursement recommendation: efgartigimod alfa injection (Vyvgart SC). *Can. J. Health Technol.* 6(1), 1–15 (2026).
33. Canada's Drug Agency. Reimbursement review: efgartigimod alfa injection (Vyvgart SC). (2025). https://canjhealthtechnol.ca/index.php/cjht/article/view/SR0894r/SR0894r
34. Ades AE, Welton NJ, Dias S, Phillippo DM, Caldwell DM. Twenty years of network meta-analysis: continuing controversies and recent developments. *Res. Synth. Methods* 15(5), 702–727 (2024).