



REVIEW

The Value of Indirect Comparisons of Systemic Biologics for Psoriasis: Interpretation of Efficacy Findings

Matthias Augustin · Christopher Schuster · Can Mert · Alexander Nast

Received: May 11, 2022 / Accepted: June 21, 2022 / Published online: July 14, 2022
© The Author(s) 2022

ABSTRACT

Introduction: It can be challenging for dermatologists to keep abreast of the growing evidence from published indirect comparisons (ICs) of treatments for psoriasis. The objective of this analysis was to summarise comparative clinical efficacy and safety findings from ICs of systemic biologics for the treatment of moderate-to-severe psoriasis and to identify factors potentially affecting efficacy outcomes and their possible implications for clinical decision making.

Methods: An umbrella review of short- and long-term efficacy and safety findings from 26

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s13555-022-00765-3>.

M. Augustin (✉)
Institute for Health Services Research in
Dermatology and Nursing, University Medical
Center Hamburg-Eppendorf, Martinistr. 52, 20246
Hamburg, Germany
e-mail: m.augustin@uke.de

C. Schuster
Eli Lilly and Company, Indianapolis, IN, USA

C. Mert
HaaPACS GmbH, Schriesheim, Germany

A. Nast
Division of Evidence-Based Medicine, Department
of Dermatology, Charité-Universitätsmedizin
Berlin, Berlin, Germany

ICs visually compared 90% improvement in Psoriasis Area Severity Index (PASI90) treatment rankings and three safety outcome rankings side by side. Pearson's correlation coefficients measured the strength of the association between each pair of ICs on the basis of identified factors that could potentially affect efficacy findings.

Results: Some consistency in short-term PASI90 efficacy rankings was observed for certain drugs, although rankings for most drugs varied by IC. Factors potentially affecting efficacy outcomes included the use of different methodologies for head-to-head comparison and statistical analyses, and variation in drugs and classes included treatment dosing and duration, outcome definitions and effect measures reported between ICs. Considerable variation in these factors was found across all 26 ICs. Comparative safety information of value to physicians was limited.

Conclusions: Substantial differences were identified between ICs in factors that could potentially affect efficacy outcomes. Treatment rankings must be interpreted alongside actual differences in IC outcomes to allow conclusions on clinical relevance. Drugs within a class cannot be considered of equal efficacy: therapies should be considered individually rather than by class.

PLAIN LANGUAGE SUMMARY

Plain language title: Interpreting indirect comparison efficacy findings in psoriasis

Psoriasis is a chronic, often life-long, inflammatory skin disease with a high disease burden and considerable quality of life impairment, particularly in moderate-to-severe psoriasis.

We, an international group of researchers, aimed to highlight the importance of fully understanding indirect comparisons (ICs) used to compare the efficacy and safety of multiple treatments for moderate-to-severe plaque psoriasis outside of head-to-head clinical trials, to effectively use their findings in clinical practice.

We visually compared side-by-side efficacy and safety results and rankings from 26 ICs of systemic biologics for psoriasis and described our findings. We also identified factors with the potential to affect the efficacy outcomes of these ICs (Fig. 2) and looked for similarities and differences between the 26 ICs in terms of these factors. Finally, given what we found, we considered how the results of these ICs should be interpreted to support physicians in choosing the treatments best suited to the needs of their patients with psoriasis.

Treatment ranking results from ICs (based on, for example, efficacy or safety outcomes) need to be interpreted together with information on actual numerical differences between treatment effects to inform the physician's decisions. When selecting the most efficacious treatment, drugs within a class cannot be considered equal, and therapies should be considered individually rather than by class. Substantial differences were identified between ICs with respect to factors that could potentially affect efficacy outcomes, highlighting the need to fully understand each individual analysis before making decisions based on its findings. Comparative safety information was limited.

Keywords: Biologics; Psoriasis; Indirect comparison; Network meta-analysis; Efficacy

Key Summary Points

What's already known about this topic?

The use of indirect comparison (IC) methodologies to compare the efficacy and safety of multiple interventions for psoriasis is growing, and methods used to perform these analyses are diversifying, making it increasingly difficult for dermatologists to keep abreast of the available data and to fully understand the perspective, methods and quality of each individual analysis.

What does this study add?

This review identified some consistency in short-term efficacy rankings for certain systemic biologic drugs for the treatment of psoriasis, although rankings for most drugs varied by IC. Factors potentially affecting efficacy outcomes varied considerably across all ICs.

In psoriasis, ICs do not yet provide sound comparative safety or long-term efficacy information of value to physicians; however, long-term efficacy data should be forthcoming from clinical trials in the near future.

What are the clinical implications of this work?

Considerable variation in factors potentially affecting efficacy outcomes across ICs means a detailed understanding of the scope and conduct of each IC is crucial prior to using its findings to inform clinical decisions.

Treatment rankings need to be interpreted alongside actual differences in outcomes to allow conclusions on clinical relevance. Drugs within a class cannot be considered equal in terms of efficacy and, therefore, should be considered individually.

INTRODUCTION

In recent years, there has been a marked increase in the number of clinical studies and ensuing publications on biologic therapies for the treatment of moderate-to-severe psoriasis. Unsurprisingly, the number of published indirect comparisons [ICs, including network meta-analyses (NMAs)], has grown in parallel [1]. As a result, it has become increasingly difficult to keep an up-to-date overview of the available evidence on the efficacy and safety of systemic biologics for psoriasis, as well as understanding the perspective and quality of that evidence.

ICs include adjusted indirect comparison (AIC) between two treatments made through a common comparator [2] and NMAs, which use a combination of indirect and direct evidence to compare more than two treatments in a single analysis [3]. Results of ICs fill an important evidence gap, as head-to-head comparisons of all available treatments are not feasible [3, 4]. However, as discussed in a previous paper in this series [1], the results of some ICs—particularly more complex NMAs—can be misleading if misinterpreted owing to a limited understanding of the research objectives and/or methods used. Further shortcomings have been described in a recent review on formal requirements of NMA for psoriasis [5]. Nevertheless, the authors revealed fair consistency of outcomes across the 27 NMAs included.

The aim of this analysis was to summarise the comparative clinical efficacy and safety findings, and resulting treatment rankings, of systemic biologics for the treatment of moderate-to-severe plaque psoriasis, as reported by identified ICs. We also aimed to identify factors potentially affecting efficacy outcomes and their possible implications for clinical decision making. Our findings should allow dermatologists to more confidently compare the efficacy of systemic biologics for psoriasis in the absence of head-to-head trials for all available treatment options and to support physicians in choosing the treatments best suited to the needs of their patients.

MATERIALS AND METHODS

A systematic literature review, which identified published ICs (including AICs and NMAs) of biologics for the treatment of moderate-to-severe psoriasis in adults (≥ 18 years) to March 2020, has been described in detail elsewhere (PROSPERO CRD42020163081) [1, 6]. The 26 analyses identified are listed in Table 1, along with descriptions of type of IC, assessment timepoint, efficacy and safety outcome(s), effect measure(s) reported and method of presentation of results/treatment ranking.

In this umbrella review of published IC results in psoriasis, short- and long-term efficacy and, where available, safety findings were summarised across the four AICs and 22 NMAs [1]. Efficacy endpoints of interest were proportions of patients achieving 75% or 90% improvement in Psoriasis Area Severity Index (PASI75 or 90) versus placebo or active comparator in the short term (induction phase) or long term (> 40 weeks). Safety outcomes of interest were serious adverse events (SAEs), number of patients with at least one adverse event (AE) and treatment discontinuation/study withdrawal due to AEs. The primary analysis was a comparison of short- and long-term PASI90 efficacy rankings across ICs. Consistencies and inconsistencies in reported results across the ICs were assessed, and factors potentially affecting outcomes were identified and considered. Safety rankings were also compared across ICs where possible.

This article is based on previously conducted studies and does not contain any new studies with human participants or animals performed by any of the authors.

Short- and Long-Term PASI90 Rankings

Most publications reported short-term PASI90 outcome data and provided treatment rankings as well as risk differences between therapies to allow for assessments of relative efficacy. Three analyses (all NMAs) included long-term PASI90 data. Available PASI90 efficacy treatment rankings for all evaluated drugs were visually compared side by side for each AIC and NMA

Table 1 Overview of adjusted indirect comparisons and network meta-analyses included in this analysis ($N = 26$) [4, 7–31]

Publication	Type of indirect comparison	Short versus long term	Efficacy and safety outcome(s)	Effect measure(s)	Method of presentation of results/treatment ranking
Reich et al. 2012 [7]	NMA	Short term (12 weeks)	Efficacy: PASI50, 75 and 90 No safety	Probability of response and relative risk (95% CrI)	League table
Lin et al. 2012 [8]	NMA	Short term (10–16 weeks)	Efficacy: PASI50, 75 and 90 No safety	Odds ratio (95% CrI)	League table and forest plot
Gupta et al. 2014 [28]	NMA	Short term (10–16 weeks)	Efficacy: PASI75 No safety	Odds ratio (95% CrI)	League table
Messori et al. 2015 [31]	NMA	Short term (12–24 weeks)	No efficacy Safety: any infectious AE, SAE	Risk difference (95% CrI)	Forest plot and ranking histogram
Signorovitch et al. 2015 [9]	NMA	Short term (8–16 weeks)	Efficacy: PASI50, 75 and 90 No safety	Posterior mean (95% CI) and rate ratio (95% CrI)	League table
Sbidian et al. 2017 [10]	NMA	Short term (12–16 weeks)	Efficacy: PASI75 and 90, PGA (0,1) Safety: AE, SAE	Risk ratio and standardised mean difference (95% CIs)	League table, forest plot and SUCRA
Jabbar-Lopez et al. 2017 [29]	NMA	Short term (12–16 weeks)	Efficacy: minimal residual activity/PASI > 90/PGA (0,1) and PASI75 Safety: d/c due to AE	Odds ratio and NNT (95% CIs)	League table and SUCRA
Gómez-García et al. 2017 [11]	NMA	Short term (10–16 weeks)	Efficacy: PASI75 and 90, IGA/PGA/sPGA (0,1) Safety: > 1 AE, infectious AE, SAE, d/c due to AE	Odds ratio (95% CI)	League table and SUCRA
Sawyer et al. 2018 [12]	NMA	Short term (10–16 weeks)	Efficacy: PASI50, 75, 90 and 100 No safety	Risk ratio (95% CrI)	League table and forest plot

Table 1 continued

Publication	Type of indirect comparison	Short versus long term	Efficacy and safety outcome(s)	Effect measure(s)	Method of presentation of results/treatment ranking
Lv et al. 2018 [13]	NMA	Short term (6–24 weeks) ^a	Efficacy: PASI50, 75 and 90, PGA (0,1) Safety: all AEs, nasopharyngitis, upper respiratory tract infection, infection, headache, d/c due to AE	Odds ratio and mean difference (95% CrI)	League table, forest plot and SUCRA
Loos et al. 2018 [14]	NMA	Short term (10–16 weeks)	Efficacy: PASI50, 75 and 90 Safety: any AE, SAE, \geq grade 3 AE, specific AE, serious infection, d/c due to AE, treatment-related death	Relative risk ratio (95% CrI)	League table
Geng et al. 2018 [25]	NMA	Long term (up to 18.3-month follow-up)	Efficacy: PASI50, 75 and 90 No safety	Odds ratio (95% CrI)	League table and rank analysis
Cameron et al. 2018 [15]	NMA	Short term (16 weeks)	Efficacy: PASI90 No safety	Risk ratio and risk difference (95% CrI)	League table, forest plot and SUCRA
Xu et al. 2019 [16]	NMA	Short term (12–16 weeks)	Efficacy: PASI50, 75, 90 and 100, PGA Safety: headache, infection, d/c due to AE	Odds ratio (95% CrI)	League table, forest plot and SUCRA
Sawyer et al. 2019 [17]	NMA	Short term (10–16 weeks)	Efficacy: PASI50, 75, 90 and 100, NNT for PASI100 No safety	Risk ratio and NNT (95% CrI)	League table
Sawyer et al. 2019 [26]	NMA	Long term (up to 52-week follow-up)	Efficacy: PASI 75, 90 and 100 No safety	Risk ratio (95% CrI)	League table, SUCRA and ranking probability plots

Table 1 continued

Publication	Type of indirect comparison	Short versus long term	Efficacy and safety outcome(s)	Effect measure(s)	Method of presentation of results/treatment ranking
Cameron et al. 2019 [18]	NMA	Short term (10–16 weeks)	Efficacy: PASI50, 75, 90 and 100, PGA/IGA (0,1) Safety: AE, SAE, d/c due to AE	Relative risk and risk difference (95% CrI)	League table, forest plot and SUCRA
Bai et al. 2019 [30]	NMA	Short term (12–16 weeks)	Efficacy: PASI75 and 100, sPGA/IGA/PGA (0,1) Safety: AE, SAE, d/c due to AE	Relative risk (95% CrI)	League table, forest plot and SUCRA
Warren et al. 2020 [4]	NMA	Short term (12 weeks)	Efficacy: PASI75, 90 and 100 No safety	Posterior mean and treatment effect relative to comparator (95% CrI)	Forest plot
Warren et al. 2020 [19]	NMA	Short term (12–16 weeks)	Efficacy: PASI75, 90 and 100 No safety	Area under the curve (95% CrI)	Forest plot, bell curve and rank analysis
Sbidian et al. 2020 [20]	NMA	Short term (8–24 weeks)	Efficacy: PASI75 and 90, PGA (0,1) Safety: AE, SAE	Risk ratio and standardised mean difference (95% CIs)	League table, forest plot and SUCRA
Armstrong et al. 2020 [21]	NMA	Short term (10–16 weeks) and long term (44–60 weeks)	Efficacy: PASI75, 90 and 100, and respective NNTs No safety	Posterior median, odds ratio and NNT (95% CrI), and response rate (95% CI)	League table and forest plot
Galván-Banqueri et al. 2013 [22]	AIC	Short term (12–24 weeks)	Efficacy: PASI50, 75 and 90 No safety	Absolute risk reduction (95% CI)	Descriptive table
Schmitt et al. 2014 [27]	Random-effects meta-analysis ^b	Short term (8–16 weeks)	Efficacy: PASI50, 75 and 90 ^c No safety	Risk difference (95% CI)	League table and forest plot

Table 1 continued

Publication	Type of indirect comparison	Short versus long term	Efficacy and safety outcome(s)	Effect measure(s)	Method of presentation of results/treatment ranking
Warren et al. 2018 [23]	MAIC	Short term (12 weeks)	Efficacy: PASI75, 90 and 100 No safety	Risk difference and odds ratio (95% CIs)	Forest plot
Papp et al. 2018 [24]	‘MAIC-like’	Short term (12 weeks)	Efficacy: PASI75, 90 and 100, symptom resolution, lesion resolution and CDC Safety: AE	Outcomes before and after propensity score weighting	Histogram and descriptive table

AE adverse event, *AIC* adjusted indirect comparison, *CDC* complete disease control, *CI* confidence interval, *CrI* credibility interval, *d/c* discontinuation, *IGA* Investigator’s Global Assessment, *MAIC* matching-adjusted indirect comparison, *NMA* network meta-analysis, *NNT* number needed to treat, *PASI50, 75, 90 and 100*, Psoriasis Area and Severity Index 50%, 75%, 90% and 100% reduction from baseline, *PGA* Physician’s Global Assessment, *SAE* serious adverse event, *sPGA* static Physician’s Global Assessment, *SUCRA* surface under the cumulative ranking curve

^aNMA also included two long-term (> 40-week) studies

^bThe Schmitt et al. 2014 [27] publication primarily described a meta-analysis; however, it also included an AIC

^cAlthough the Schmitt et al. 2014 [27] publication indicated that PASI90 outcomes were considered, no comparative PASI90 outcome data were reported in the manuscript

reporting short- or long-term PASI90 response data ($N = 21$) [4, 7–26].

Factors Potentially Affecting Efficacy Outcomes

Factors that could potentially affect the efficacy findings of ICs were identified by the authors through review of the literature concerning this topic. For example, different sources of sponsorship, IC methods (e.g. NMA versus AIC) or statistical approaches (e.g. Bayesian versus frequentist), disparate characteristics of the studies included, variety of outcomes considered (e.g. PASI75 versus 90) and/or results presentation methods used. Each analysis was then assigned yes/no answers to a series of questions based on the impact of the identified factors. A dataset was generated where each column defined one IC and each row defined an identified factor. For

each IC, a value of 1 was assigned in the respective row for the presence of a factor that could have impacted efficacy findings (i.e. a yes answer), while a value 0 was assigned for the absence of this factor (i.e. a no answer).

Pearson’s correlation coefficient (r , from -1 to $+1$) was used to assess the similarity between each pair of ICs (both AICs and NMAs) with respect to the factors included in the analysis. The magnitude of the correlation coefficient describes the strength of the similarity or dissimilarity. A correlation value of 1 refers to perfect similarity, whereas -1 refers to perfect dissimilarity. The Pearson’s correlation coefficient analysis was run in R version 4.0.5 (<https://www.rstudio.com>).

Short- and Long-Term Clinical Efficacy Outcomes

We graphically presented clinical efficacy results for currently licensed biologics from identified ICs alongside each other for comparison wherever possible.

Short-term PASI90 outcomes versus placebo, reported using comparable effect measures and organised by treatment class, were compared for 11 NMAs [7, 8, 10–17, 20]. Additional short-term PASI90 outcomes were available from five further NMAs; however, diversity in reported effect measure(s) and/or treatment comparators (i.e. placebo or actives) between analyses made side-by-side comparisons of these findings more challenging [4, 9, 18, 19, 21]. Short-term PASI90 outcomes were available for three AICs, although once again methodological differences between analyses hindered side-by-side comparisons of findings [22–24]. One AIC reported only short-term PASI75 outcomes [27]. Similarly, three NMAs did not report PASI90 outcomes; therefore, short-term PASI75 outcomes versus placebo, organised by treatment class, were presented for these analyses [28–30].

Long-term PASI90 outcomes were compared from three NMAs, organised by treatment class [21, 25, 26]. Although assessment timepoints for these three analyses varied considerably, treatment comparators and reported effect measure(s) were consistent enough to allow side-by-side graphic presentations of long-term efficacy findings for two of the three NMAs.

One NMA reported safety data only, so was excluded from the entire clinical efficacy analysis [31].

Clinical Safety Rankings

Limited safety data were included in the four AIC and 22 NMA publications identified, with no long-term safety data reported (Table 1). Available short-term safety rankings for all evaluated drugs were visually compared for each NMA reporting safety outcomes of interest; no AICs reported safety rankings. Rankings were presented side by side for each of short-term SAE (five NMAs) [10, 18, 20, 30, 31], the

number of patients with at least one AE (six NMAs) [10, 11, 13, 18, 20, 30] and treatment discontinuation/study withdrawal due to AEs (five NMAs) [13, 16, 18, 29, 30].

RESULTS

Short- and Long-Term PASI90 Rankings

Figure 1 shows the short- or long-term PASI90 treatment rankings reported in each AIC and NMA included in this umbrella review. This figure illustrates many of the challenges associated with comparing efficacy findings for individual drugs, or even classes of drugs, across different ICs: different treatments were included in each analysis (and some ICs included the same treatments but different trials); dosing information was reported for some ICs but not for others (and some ICs pooled doses whereas others analysed them separately); ranking methods differed, introducing uncertainty; and data collection timepoints varied, even when outcome measures were consistent.

Despite these challenges, some consistency in short-term PASI90 efficacy rankings was observed for certain drugs, although rankings for most drugs varied by AIC or NMA, making a detailed understanding of the scope and conduct of each AIC or NMA crucial. When included in ICs, the newer drug classes, interleukin (IL)-17 and IL-23 antagonists, with the notable exception of tildrakizumab, generally ranked more highly overall than the older drugs, tumour necrosis factor (TNF) and IL-12/23 antagonists (as illustrated by green and purple groupings to the left of Fig. 1). However, infliximab ranked more highly than some or all of these newer treatments in several ICs and was usually the highest-ranking TNF antagonist across ICs, illustrating the importance of looking at the results for each drug individually, not grouped together by class. Furthermore, IL-17 antagonists appeared to be ranked consistently highly with respect to PASI90 outcomes where they were included in analyses. Of note, brodalumab and ixekizumab usually outranked secukinumab when all three IL-17 antagonists were included in a single NMA.

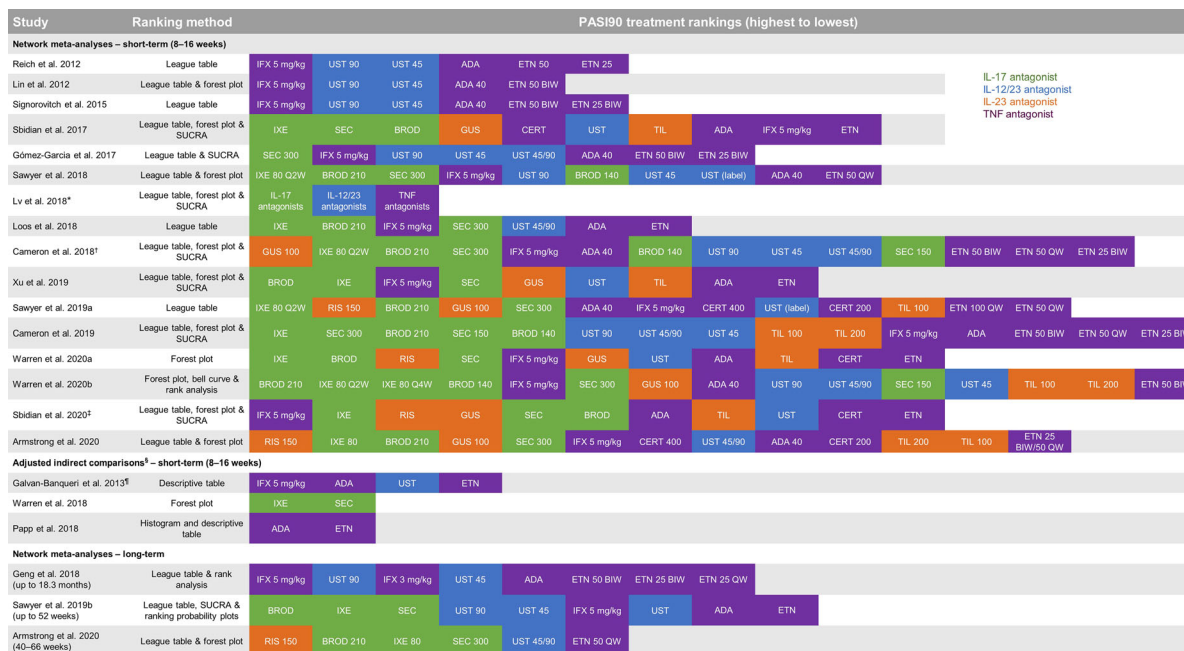


Fig. 1 Treatment rankings for adjusted indirect comparisons and network meta-analyses reporting PASI90 data (N = 21) [4, 7–26]. Drug dosages were not always reported in the publications considered. Doses are in milligrams, unless indicated otherwise, and as reported in individual AICs or NMAs. PASI90 data from individual NMAs are not reported for alefacept, bimekizumab, briakinumab, efalizumab, itolizumab, ponesimod or tofacitinib, as these drugs have either been removed from the market or have not yet received market approval for this indication in Europe, the USA or Japan. *NMA also included two long-term (> 40-week) studies. †Ranking in adjusted analyses; top five rankings in unadjusted analyses IXE 80 Q2W, IFX 5 mg/kg, BROD 210, SEC 300, GUS 100. ‡Short-term efficacy assessed at 8–24 weeks. §Ranking

based on analyses themselves, no additional ranking analysis undertaken. ¶Short-term efficacy assessed at 12–24 weeks. ADA adalimumab, AIC adjusted indirect comparison, AUC area under curve, BIW twice weekly, BROD brodalumab, CERT certolizumab, EMA European Medicines Agency, ETN etanercept, FDA Food and Drug Administration, GUS guselkumab, IFX infliximab, IL interleukin, IXE ixekizumab, NMA network meta-analysis, PASI90, ≥ 90% improvement from baseline in Psoriasis Area Severity Index, QW once weekly, Q2W/Q4W every 2/4 weeks, RIS risankizumab, SEC secukinumab, SUCRA surface under the cumulative ranking, TIL tildrakizumab, TNF tumour necrosis factor, UST ustekinumab

Factors Potentially Affecting Efficacy Outcomes

The factors considered by the authors as potentially affecting the efficacy findings of AICs and NMAs are described in Fig. 2. These included not only the use of different methodologies for head-to-head comparison and statistical analyses, but also the variation between AICs and NMAs with respect to study designs, drugs and classes included, treatment dosing and duration, as well as outcome definitions and effect measures reported.

Figure 3 illustrates the considerable level of variation in these factors across all 26 ICs. Minimal associations were found between the four AICs; Pearson’s correlation coefficients ranged from –0.2 to 0.3 for all AIC pairwise comparisons. The range of correlation coefficients for all NMA pairwise comparisons was broader (–0.5 to 0.9), suggesting that some NMAs exhibited striking differences while others were very similar, although many showed no or nominal correlation. Some older NMAs published in the same year (2012) [7, 8] showed strong associations (correlation coefficient 0.6), but those published in 2020 [4, 19–21] did not

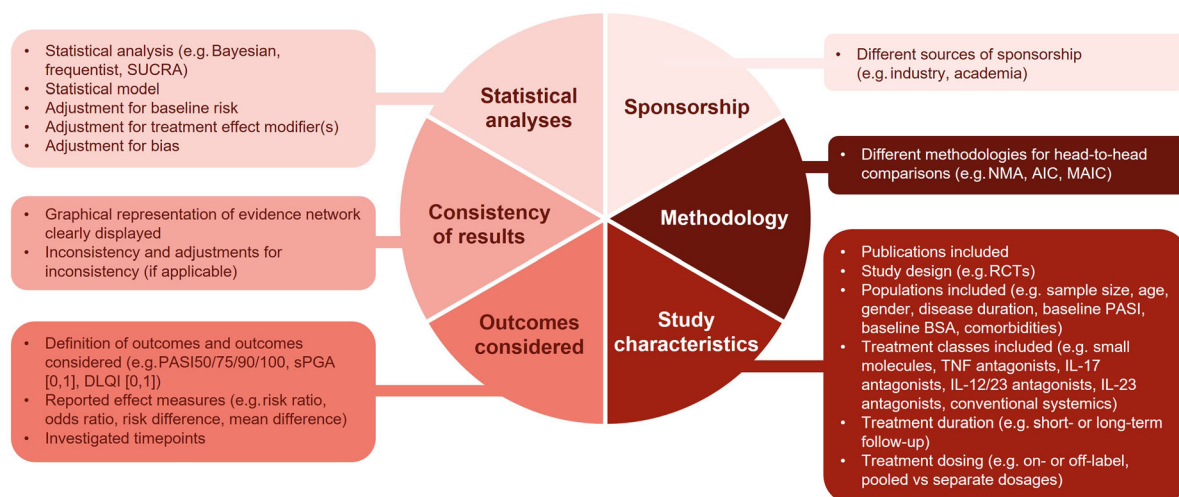


Fig. 2 Factors potentially affecting adjusted indirect comparison and network meta-analysis outcomes identified through review of the literature. *AIC* adjusted indirect comparison, *BSA* body surface area, *DLQI* Dermatology Life Quality Index, *IL* interleukin, *MAIC* matching-adjusted indirect comparison, *NMA* network meta-

analysis, *PASI* Psoriasis Area Severity Index, *PASI50/75/90/100* $\geq 50\%/ \geq 75\%/ \geq 90\%/100\%$ improvement from baseline in Psoriasis Area Severity Index, *RCT* randomised controlled trial, *sPGA* static Physicians Global Assessment, *SUCRA* Surface Under the Cumulative Ranking, *TNF* tumour necrosis factor

consistently show the same pattern of similarities (from -0.1 to 0.9). Perhaps unsurprisingly, NMAs authored by the same teams, but with slightly different purposes, appeared to be similar in terms of the factors considered (correlation coefficients from 0.6 to 0.9).

Short- and Long-Term Clinical Efficacy Outcomes

The figures presenting short- or long-term efficacy outcomes from multiple ICs side by side, as described in the Methods section, are included in Supplementary Material. Overall, the estimated efficacy of individual drugs varied across AICs and NMAs. Efficacy differences were generally higher within a drug class than across classes.

Clinical Safety Rankings

Safety rankings, based on the limited short-term safety data available across the NMAs, are presented in Fig. 4.

Little knowledge can be gleaned from these safety findings owing to the sparse data and variations in outcomes, endpoint definitions and collection timepoints within individual analyses. IL-23 antagonists appeared to be consistently more highly ranked than IL-17 antagonists with respect to the short-term number of patients with at least one AE outcome (Fig. 4b), although this pattern was not visible in the SAE or treatment discontinuation/study withdrawal owing to AE safety rankings (Fig. 4a, c, respectively). However, it is important to note that the AE outcome definitions varied from NMA to NMA, with some analyses evaluating ‘at least one AE’ and others describing ‘incidence of AE’, or alternative definitions. Also of note, the IL-12/23 antagonist, ustekinumab, seemed to retain a steady high-to-mid-ranking position across all three safety outcome rankings.

DISCUSSION

We aimed to summarise comparative clinical efficacy and safety findings from ICs of systemic biologics for the treatment of moderate-to-

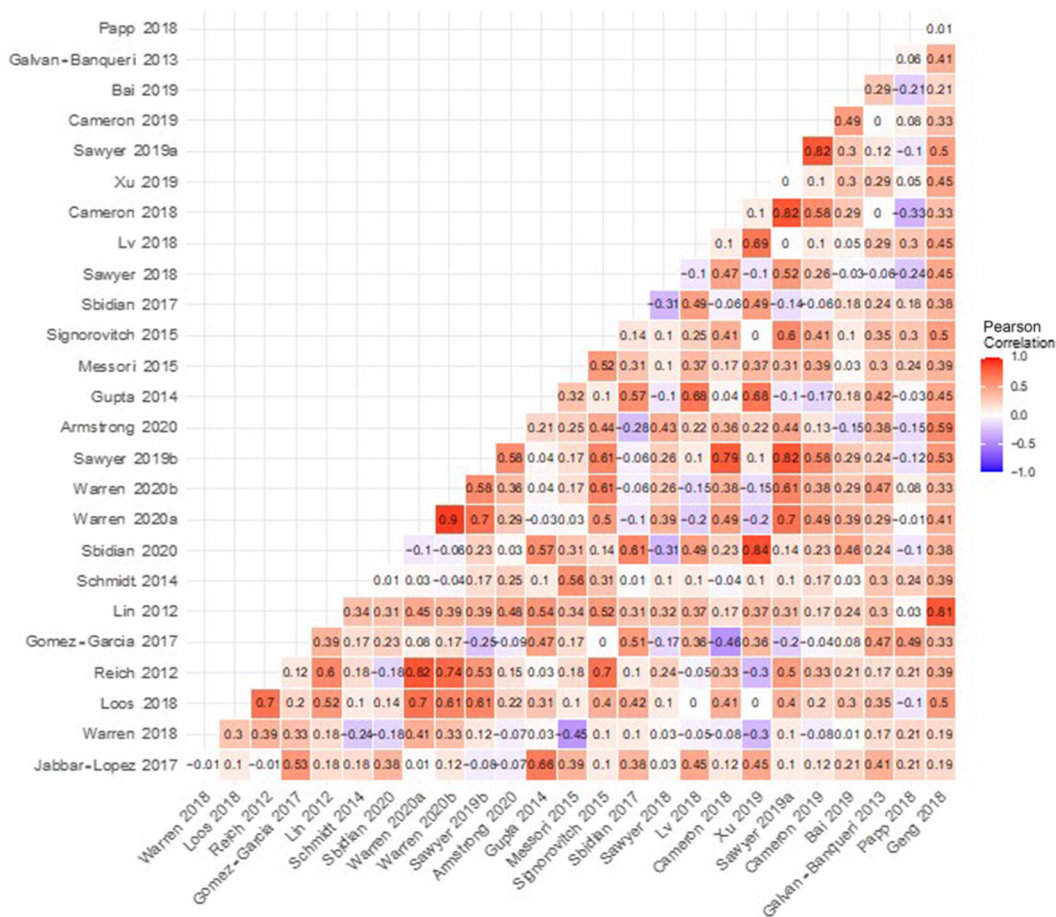


Fig. 3 Pearson’s correlation coefficients calculated to assess the levels of similarity or difference between each pair of indirect comparisons (adjusted indirect comparisons and network meta-analyses) included in this analysis ($N = 26$) with respect to factors potentially affecting outcomes. Factors were identified by the authors and outlined in Fig. 2. A Pearson’s correlation coefficient value of 1 refers to perfect similarity (dark red) where -1 refers to perfect dissimilarity (dark purple). Perfect similarity or

dissimilarity between the ICs indicated the absence or presence of factors that could affect efficacy outcomes between each pair of analyses. The magnitude of the correlation coefficient describes the strength of the similarity or dissimilarity, with darker colours indicating values closer to 1 or -1 and lighter colours indicating values closer to 0. The coefficients do not have any clinical meaning per se. *IC* indirect comparison

severe psoriasis and identify factors potentially affecting efficacy outcomes and their possible implications for clinical decision making. Our umbrella review of previously identified ICs [1] found broad patterns across short- and long-term PASI90 efficacy rankings by drug; however, the exact rankings of individual drugs were rarely replicated in the different ICs. Factors with the potential to affect IC efficacy findings varied considerably across analyses.

Of note, updates to ICs can show strong changes in treatment rankings, as newly included trials can reverse a previously determined order. For example, Sbidian et al. [10, 20] performed an NMA in 2016 and updated it in 2020 using the same ranking method. The 2016 NMA included no direct comparisons and used only indirect data, whereas the 2020 NMA was based on five trials providing data for the target outcome and time interval. In 2016, infliximab ranked second-to-last in terms of clinical

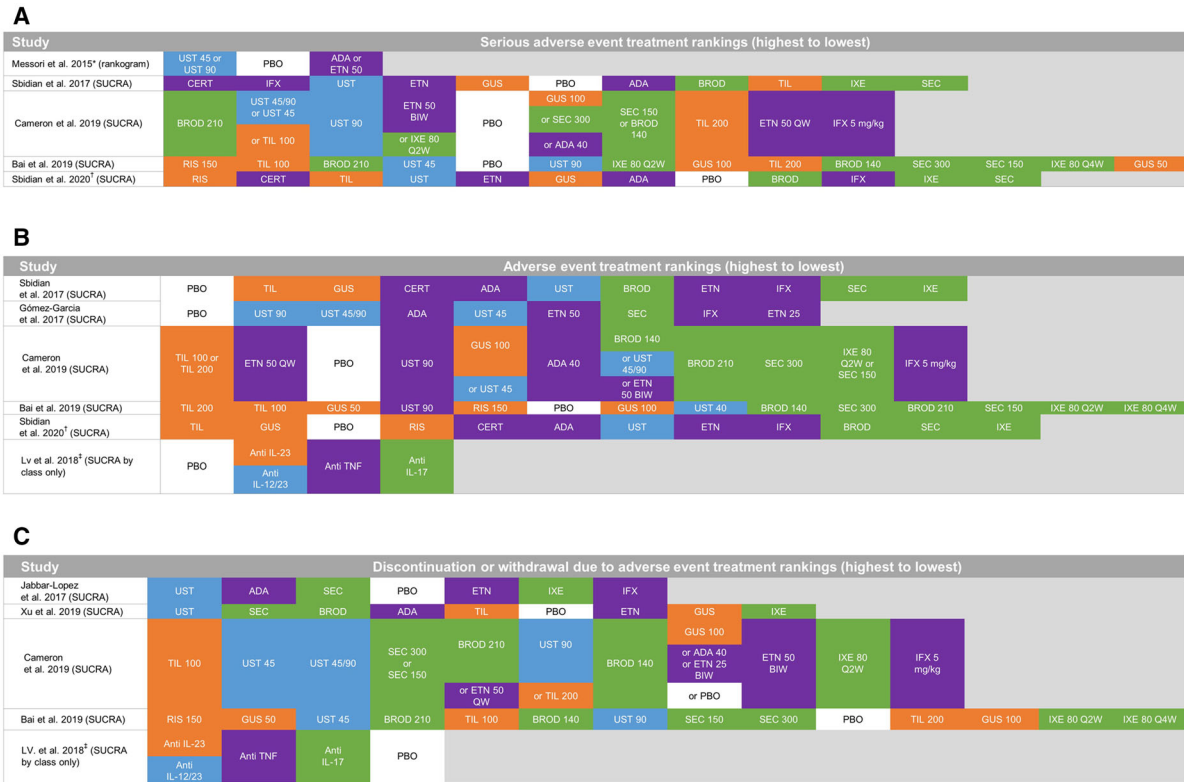


Fig. 4 Safety rankings: results from nine network meta-analyses reporting short-term (weeks 8–16) serious adverse events, number of patients with at least one adverse event, or treatment discontinuation/study withdrawal due to adverse event outcomes [10, 11, 13, 16, 18, 20, 29–31]. The highest ranking indicates the best safety profile. Some treatments were ranked equally, as indicated by ‘or’ in a coloured box. Doses are in milligrams, unless indicated otherwise, and as reported in individual NMAs. Not all NMAs reported drug doses. Safety data from individual NMAs are not reported for alefacept, bimekizumab, briakinumab, efalizumab, itolizumab, ponesimod or tofacitinib, as these drugs have either been removed from the market or have not yet received market approval for this indication in Europe, the USA or Japan. *Short-term safety

assessed at 12–24 weeks. †Short-term safety assessed at 8–24 weeks. ‡Short-term safety assessed at 6–24 weeks and NMA included two long-term (> 40-week) studies. ADA adalimumab, BIW twice weekly, BROD brodalumab, CERT certolizumab, ETN etanercept, GUS guselkumab, IFX infliximab, IL interleukin, IXE ixekizumab, PBO placebo, Q2W/Q4W every 2/4 weeks, QW once weekly, RIS risankizumab, SEC secukinumab, SUCRA surface under the cumulative ranking, TIL tildrakizumab, TNF tumour necrosis factor, UST ustekinumab. **a** Serious adverse events [10, 18, 20, 30, 31]. **b** Number of patients with at least one adverse event [10, 11, 13, 18, 20, 30]. **c** Treatment discontinuation/study withdrawal due to adverse events [13, 16, 18, 29, 30]

efficacy (PASI90) among the ten biologics compared using the surface under the cumulative ranking (SUCRA) method (low certainty of evidence) [10]; however, in 2020, infliximab ranked highest (above newer drugs such as ixekizumab, risankizumab and guselkumab) (moderate certainty of evidence), demonstrating that estimates based on low certainty of evidence can easily be reversed by new data

[20]. These types of inconsistencies were observed in a number of cases, highlighting the importance of considering how or if the methodology used to undertake an NMA or AIC, the studies included, the inconsistencies identified and/or adjusted for, the statistical analyses undertaken—among other factors—may have affected the results. These findings contradict a prior publication, which showed

markedly better concordances between studies [5].

Two Warren et al. NMA publications [4, 19] provide an interesting example of how estimates of effect may differ even when there are only minor differences in aim and design between two analyses. These papers, both published in 2020, reported similar NMAs conducted from different perspectives: one clinical and one statistical. Although both NMAs assessed three IL-17 antagonists (brodalumab, ixekizumab and secukinumab) at a similar point in time, treatment doses were reported and handled differently, and different effect measures and ranking methods were used. As a result, each NMA reported different treatment rankings.

Our PASI90 efficacy ranking analysis also illustrated the importance of considering treatments individually, not grouped together by class. Tildrakizumab consistently ranked lower in the short term than other IL-23 antagonists, often even lower than older IL-12/23 and TNF antagonists, whereas guselkumab and risankizumab generally ranked highly, along with the IL-17 antagonists. In this case, NMA efficacy findings may have been affected by the choice of investigated timepoints included in the analyses, as tildrakizumab is known to have a longer time to treatment effect compared with other IL-23 antagonists [32–35], although tildrakizumab was not included in any long-term study. Regardless, were the IL-23 antagonists considered as a class, not individually, the results for the class would be skewed by these rankings [20].

The ranking method used to summarise the results of an NMA (or AIC) may also have an impact on the results and the risk of misinterpretation could have implications for treatment decisions that may not be in the best interests of patients. Ranking approaches should be viewed with caution because they are often based on a single outcome measure (e.g. PASI90 in this analysis), although there are typically several outcomes measures of interest, and some are presented without an accompanying measure of uncertainty [36, 37]. Furthermore, rankings may not consider the magnitude of differences in effects between treatments or capture the

possibility that chance may explain differences between treatments [37]. Moreover, without understanding the level of certainty of the evidence on which the rankings are based, one cannot interpret the rankings with much confidence.

Of the factors identified as potentially affecting the efficacy outcomes of ICs, some appeared to be similar across publications, such as the study designs, the analysis of patient populations (i.e. adult patients with moderate-to-severe psoriasis), the efficacy outcome measures (e.g. PASI75, 90 and 100) and the meta-analytic methods used. However, other factors varied considerably; for example, the sponsorship source, the selection of treatments or treatment classes for inclusion, the reporting and handling of doses, the effect measure(s) reported and the methods used to present results and treatment rankings [1]. Our assessment of the overall similarities and differences with respect to these factors between the identified ICs suggested more differences than similarities between the analyses, with generally low (-0.5 to $+0.5$) Pearson's correlation coefficients. These findings reflect the results of our PASI90 efficacy ranking analysis, which showed little consistency in individual treatment rankings between analyses. Interestingly, even some NMAs authored by similar teams using similar designs and methodologies, as reflected by higher positive Pearson's correlation coefficients, reported inconsistent treatment rankings. This emphasises the importance of understanding all these factors prior to interpreting AIC or NMA results.

Few ICs were identified to inform our long-term PASI90 efficacy treatment ranking analysis at > 40 weeks; however, the limited long-term efficacy rankings broadly mirrored the trends observed in the short-term treatment rankings described previously. Historically, crossover or re-randomisation study designs have made NMAs of long-term data challenging, and use of the placebo response-carried-forward approach is associated with significant uncertainties. Furthermore, traditional pairwise meta-analysis can only provide a summary of available direct evidence without comparisons between treatments for one intervention relative to another,

making it insufficient to guide practical decision making. Nonetheless, the long-term efficacy data from head-to-head studies with a primary endpoint at > 40 weeks in moderate-to-severe psoriasis required to undertake long-term ICs are now slowly emerging; therefore, future NMAs should be able to evaluate efficacy data beyond the induction phase.

Unfortunately, we were unable to learn much from the identified ICs with regard to the comparative safety of licensed biologics for the treatment of moderate-to-severe psoriasis. It would appear that NMAs in psoriasis to date have not been successful in providing sound comparative safety information of value to physicians to inform clinical decision making. This is likely due to, at least in part, the high level of variability we observed in safety parameters, endpoint definitions and collection timepoints across clinical trials for different treatments. It remains to be seen whether attempts to harmonise safety data capture across clinical trials and real-world treatment use can provide sufficiently consistent data to inform robust NMAs of safety outcomes.

Limitations

This analysis is likely to become outdated quickly, given the expected publication activity in this rapidly advancing disease area, particularly with respect to long-term data. With regard to our methods, the limitations of the NICE TSD7 (National Institute for Health and Care Excellence Technical Support Document 7) 'Evidence synthesis of treatment efficacy in decision making: a reviewer's checklist' [38] and AMSTAR 2 (A Measurement Tool to Assess Systematic Reviews version 2) [39] used in the analyses reported in the previous paper in this series [1] also apply. Furthermore, we believe this is the first time a Pearson's correlation coefficient analysis has been applied to studies rather than patients; hence, further research on the pros and cons of this method are required. A statistical analysis of all identified IC findings is not yet possible; therefore, we conducted a simple side-by-side comparison of the IC results, although a lack of consensus on reported safety

parameters limited our comparison of the safety results.

CONCLUSIONS

Current ICs, particularly NMAs, provide valuable indirect evidence of the short-term efficacy of available systemic biologic treatment options for moderate-to-severe psoriasis. However, substantial differences were identified between AICs and NMAs with respect to factors that could potentially affect efficacy outcomes. Treatment rankings need to be interpreted alongside actual differences in IC outcomes to allow conclusions on clinical relevance. When selecting the most efficacious treatment, drugs within a class cannot be considered equal, and therapies should be considered individually rather than by class. Current ICs provide few safety analyses, which have to be interpreted with caution owing to low numbers of patients and events, different outcome measures and varying definitions. Furthermore, most ICs to date have analysed short-term data, underlining the need for longer-term analyses to understand the comparative long-term efficacy of available treatment options.

ACKNOWLEDGEMENTS

Funding. This work and the Rapid Service Fee were supported by Eli Lilly and Company.

Medical Writing, Editorial and Other Assistance. The authors would like to acknowledge Daniel Saure (Eli Lilly and Company) for his statistical support, and Sue Williamson and Caroline Spencer (Rx Communications, Mold, UK) for medical writing assistance with the preparation of this manuscript, funded by Eli Lilly and Company.

Author Contributions. CS and AN contributed to the conception and design of the work. Data analysis was performed by CM. All authors were involved in interpreting the data, critically revising the article for important

intellectual content and read and approved the final manuscript.

Disclosures. Matthias Augustin has received research grants from AbbVie, Almirall, Beiersdorf, Eli Lilly, Galderma, Incyte, LEO, Menlo, MSD, Novartis, Pfizer, Regeneron, Sanofi-Genzyme and Trevi; served as a consultant for AbbVie, Almirall, Beiersdorf, Eli Lilly, Galderma, Incyte, LEO, Menlo, MSD, Novartis, Pfizer, Regeneron, Sanofi-Genzyme and Trevi; and as a lecturer for AbbVie, Almirall, Beiersdorf, Eli Lilly, Galderma, Incyte, LEO, Menlo, MSD, Novartis, Pfizer, Regeneron, Sanofi-Genzyme and Trevi. Christopher Schuster is an employee of Eli Lilly and Company. Can Mert is a contractor of HaaPACs GmbH and conducted statistical analysis for this project on behalf of Eli Lilly and Company. Alexander Nast declares no conflicts of interest.

Compliance with Ethics Guidelines. This article is based on previously conducted studies and does not contain any new studies with human participants or animals performed by any of the authors.

Data Availability. The dataset supporting the conclusions of this article are included within the article (and its Supplementary Material).

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence,

visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Nast A, Dressler C, Schuster C, et al. Methods used for indirect comparisons of systemic treatments for psoriasis. A systematic review. *Skin Health Dis.* 2022. <https://doi.org/10.1002/ski2.112>.
2. 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.* 1997;50:683–91.
3. Kim H, Gurrin L, Ademi Z, Liew D. Overview of methods for comparing the efficacies of drugs in the absence of head-to-head clinical trial data. *Br J Clin Pharmacol.* 2014;77:116–21.
4. Warren RB, See K, Burge R, et al. Rapid response of biologic treatments of moderate-to-severe plaque psoriasis: a comprehensive investigation using Bayesian and Frequentist Network meta-analyses. *Dermatol Ther (Heidelb).* 2020;10:73–86.
5. Augustin M, Valencia López M, Reich K. Network meta-analyses in psoriasis: overview and critical discussion. *J Eur Acad Dermatol Venereol.* 2021;35:2367–76.
6. National Institute for Health Research. International prospective register of systematic reviews (PROSPERO). Comprehensive literature review of network meta-analyses in psoriasis (CRD42020163081). 1 August 2020. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=163081. Accessed 17 Mar 2022.
7. Reich K, Burden AD, Eaton JN, Hawkins NS. Efficacy of biologics in the treatment of moderate to severe psoriasis: a network meta-analysis of randomized controlled trials. *Br J Dermatol.* 2012;166:179–88.
8. Lin VW, Ringold S, Devine EB. Comparison of ustekinumab with other biological agents for the treatment of moderate to severe plaque psoriasis: a Bayesian network meta-analysis. *Arch Dermatol.* 2012;148:1403–10.
9. Signorovitch JE, Betts KA, Yan YS, et al. Comparative efficacy of biological treatments for moderate-to-severe psoriasis: a network meta-analysis adjusting for cross-trial differences in reference arm response. *Br J Dermatol.* 2015;172:504–12.
10. Sbidian E, Chaimani A, Garcia-Doval I, et al. Systemic pharmacological treatments for chronic

- plaque psoriasis: a network meta-analysis. *Cochrane Database Syst Rev.* 2017;12:CD011535.
11. Gómez-García F, Epstein D, Isla-Tejera B, et al. Short-term efficacy and safety of new biological agents targeting the interleukin-23-T helper 17 pathway for moderate-to-severe plaque psoriasis: a systematic review and network meta-analysis. *Br J Dermatol.* 2017;176:594–603.
 12. Sawyer L, Fotheringham I, Wright E, et al. The comparative efficacy of brodalumab in patients with moderate-to-severe psoriasis: a systematic literature review and network meta-analysis. *J Dermatol Treat.* 2018;29:557–68.
 13. Lv J, Zhou D, Wang Y, et al. Quantitative evaluation to efficacy and safety of therapies for psoriasis: a network meta-analysis. *Mol Pain.* 2018;14:1744806918762205.
 14. Loos AM, Liu S, Segel C, et al. Comparative effectiveness of targeted immunomodulators for the treatment of moderate-to-severe plaque psoriasis: a systematic review and network meta-analysis. *J Am Acad Dermatol.* 2018;79:135-44.e7.
 15. Cameron C, Hutton B, Druchok C, et al. Importance of assessing and adjusting for cross-study heterogeneity in network meta-analysis: a case study of psoriasis. *J Comp Eff Res.* 2018;7:1037–51.
 16. Xu G, Xia M, Jiang C, et al. Comparative efficacy and safety of thirteen biologic therapies for patients with moderate or severe psoriasis: a network meta-analysis. *J Pharmacol Sci.* 2019;139:289–303.
 17. Sawyer LM, Malottki K, Sabry-Grant C, et al. Assessing the relative efficacy of interleukin-17 and interleukin-23 targeted treatments for moderate-to-severe plaque psoriasis: a systematic review and network meta-analysis of PASI response. *PLoS ONE.* 2019;14: e0220868.
 18. Cameron C, Druchok C, Hutton B. Guselkumab for the treatment of moderate-to-severe plaque psoriasis during induction phase: a systematic review and network meta-analysis. *J Psoriasis Psoriatic Arthritis.* 2019;4:81–92.
 19. Warren RB, Gooderham M, Burge R, et al. Comparison of cumulative clinical benefits of biologics for the treatment of psoriasis over 16 weeks results from a network meta-analysis. *J Am Acad Dermatol.* 2020;82:1138–49.
 20. Sbidian E, Chaimani A, Afach S, et al. Systemic pharmacological treatments for chronic plaque psoriasis: a network meta-analysis. *Cochrane Database Syst Rev.* 2020;1:CD011535.
 21. Armstrong AW, Puig L, Joshi A, et al. Comparison of biologics and oral treatments for plaque psoriasis: a meta-analysis. *JAMA Dermatol.* 2020;156:258–69.
 22. Galván-Banqueri M, Marín Gil R, Santos Ramos B, Bautista Paloma FJ. Biological treatments for moderate-to-severe psoriasis: indirect comparison. *J Clin Pharm Ther.* 2013;38:121–30.
 23. Warren RB, Brnabic A, Saure D, et al. Matching-adjusted indirect comparison of efficacy in patients with moderate-to-severe plaque psoriasis treated with ixekizumab vs. secukinumab. *Br J Dermatol.* 2018;178:1064–71.
 24. Papp KA, Yang M, Sundaram M, et al. Comparison of adalimumab and etanercept for the treatment of moderate to severe psoriasis: an indirect comparison using individual patient data from randomized trials. *Value Health.* 2018;21:1–8.
 25. Geng W, Zhao J, Fu J, et al. Efficacy of several biological therapies for treating moderate to severe psoriasis: a network meta-analysis. *Exp Ther Med.* 2018;16:5085–95.
 26. Sawyer LM, Cornic L, Levin LÅ, et al. Long-term efficacy of novel therapies in moderate-to-severe plaque psoriasis: a systematic review and network meta-analysis of PASI response. *J Eur Acad Dermatol Venereol.* 2019;33:355–66.
 27. Schmitt J, Rosumeck S, Thomaschewski G, et al. Efficacy and safety of systemic treatments for moderate-to-severe psoriasis: meta-analysis of randomized controlled trials. *Br J Dermatol.* 2014;170:274–303.
 28. Gupta AK, Daigle D, Lyons DCA. Network meta-analysis of treatments for chronic plaque psoriasis in Canada. *J Cutan Med Surg.* 2014;18:371–8.
 29. Jabbar-Lopez ZK, Yiu ZZN, Ward V, et al. Quantitative evaluation of biologic therapy options for psoriasis: a systematic review and network meta-analysis. *J Investig Dermatol.* 2017;137:1646–54.
 30. Bai F, Li GG, Liu Q, et al. Short-term efficacy and safety of IL-17, IL-12/23, and IL-23 inhibitors brodalumab, secukinumab, ixekizumab, ustekinumab, guselkumab, tildrakizumab, and risankizumab for the treatment of moderate to severe plaque psoriasis: a systematic review and network meta-analysis of randomized controlled trials. *J Immunol Res.* 2019;2019:2546161.
 31. Messori A, Trippoli S, Fadda V, et al. Subcutaneous biological treatments for moderate to severe psoriasis: interpreting safety data by network meta-analysis. *Drugs Real World Outcomes.* 2015;2:23–7.

32. Reich K, Papp KA, Blauvelt A, et al. Tildrakizumab versus placebo or etanercept for chronic plaque psoriasis (reSURFACE 1 and reSURFACE 2): results from two randomised controlled, phase 3 trials. *Lancet*. 2017;390:276–88.
33. Blauvelt A, Papp KA, Griffiths CE, et al. Efficacy and safety of guselkumab, an anti-interleukin-23 monoclonal antibody, compared with adalimumab for the continuous treatment of patients with moderate to severe psoriasis: results from the phase III, double-blinded, placebo- and active comparator-controlled VOYAGE 1 trial. *J Am Acad Dermatol*. 2017;76:405–17.
34. Reich K, Armstrong AW, Foley P, et al. Efficacy and safety of guselkumab, an anti-interleukin-23 monoclonal antibody, compared with adalimumab for the treatment of patients with moderate to severe psoriasis with randomized withdrawal and retreatment: results from the phase III, double-blind, placebo- and active comparator-controlled VOYAGE 2 trial. *J Am Acad Dermatol*. 2017;76:418–31.
35. Gordon KB, Strober B, Lebwohl M, et al. Efficacy and safety of risankizumab in moderate-to-severe plaque psoriasis (UltIMMa-1 and UltIMMa-2): results from two double-blind, randomised, placebo-controlled and ustekinumab-controlled phase 3 trials. *Lancet*. 2018;392:650–61.
36. Veroniki AA, Straus SE, Rucker G, Tricco AC. Is providing uncertainty intervals in treatment ranking helpful in a network meta-analysis? *J Clin Epidemiol*. 2018;100:122–9.
37. Mbuagbaw L, Rochweg B, Jaeschke R, et al. Approaches to interpreting and choosing the best treatments in network meta-analysis. *Syst Rev*. 2017;6:79.
38. National Institute for Health and Care Excellence (NICE) Decision Support Unit (DSU) Technical Support Document 7: Evidence synthesis of treatment efficacy in decision making: a reviewer's checklist. 2012. http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD7-reviewer-checklist.final_.08.05.12.pdf. Accessed 17 Mar 2022.
39. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008.