



Assessing the Quality and Coherence of Network Meta-Analyses of Biologics in Plaque Psoriasis: What Does All This Evidence Synthesis Tell Us?

Emily Wright · Najeeda Yasmeen · Kinga Malottki · Laura M. Sawyer ·
Emma Borg · Carsten Schwenke · Richard B. Warren

Received: September 24, 2020 / Published online: December 22, 2020
© The Author(s) 2020

ABSTRACT

Introduction: A range of treatments are available for moderate-to-severe psoriasis; however, there remains a paucity of direct comparisons of these in head-to-head trials. Network meta-analyses (NMA) allow comparisons of these to support clinical decision making. This systematic literature review assesses the methodological quality of NMAs available for moderate-to-severe psoriasis and compares their methods and results. Their validity and applicability for current practice is also assessed.

Methods: A systematic review of published NMAs of at least two biologics for moderate-to-severe psoriasis was undertaken. Embase,

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13555-020-00463-y>) contains supplementary material, which is available to authorized users.

E. Wright · N. Yasmeen · K. Malottki ·
L. M. Sawyer (✉)
Symmetron Limited, London, UK
e-mail: lsawyer@symmetron.net

E. Borg
LEO Pharma A/S, Ballerup, Denmark

C. Schwenke
SCO:SSiS, Berlin, Germany

R. B. Warren
Dermatology Centre, Salford Royal NHS Foundation
Trust, University of Manchester, NIHR Biomedical
Research Centre, Manchester, UK

MEDLINE, MEDLINE In-Process, and the Cochrane Library were last searched on 19 February 2020. The quality of NMAs was assessed using the International Society of Pharmacoeconomics and Outcomes Research (ISPOR) criteria. NMA methodology, funding, and results were compared and differences in results explored.

Results: Twenty-five analyses evaluating up to 19 different treatments at 8–24 weeks, and two analyses at 1 year, were included. Psoriasis Area Severity Index (PASI) response was assessed in 23, facilitating comparisons between NMAs. All NMAs met at least half of the ISPOR criteria. The major limitations were explaining the rationale for methodology, exploring effect modifiers, and consistency between direct and indirect estimates. The analyses differed in model type (Bayesian or frequentist), analysis of PASI response (binomial or multinomial), and analysis of different treatment doses (separate or pooled). PASI results were broadly similar, except for the Cochrane Collaboration NMA which provided lower estimates of treatment efficacy versus placebo. This analysis differed methodologically from others, including pooling data for different doses.

Conclusions: Based on PASI 90 at induction, the majority of recent NMAs came to similar conclusions: interleukin (IL) 17 inhibitors (brodalumab, ixekizumab, secukinumab), IL-23 inhibitors (guselkumab and risankizumab) and infliximab were most efficacious, supporting

the validity of NMAs in this clinical area. Decisions should be made using high-quality, up-to-date NMAs with assumptions relevant to clinical practice.

Keywords: Biologic therapy; Network meta-analysis; Psoriasis; Systematic review

Key Summary Points

Why carry out this study?

Clinicians and healthcare payers have a multitude of options for treating moderate-to-severe plaque psoriasis, and network meta-analysis (NMA), a method by which multiple interventions can be compared simultaneously in a single analysis, has been used widely to support decision making based on the best available evidence of efficacy and safety.

Despite the widespread application of NMA to synthesize randomized evidence in psoriasis, skepticism around its use and mistrust of its results persist.

To address these concerns, we performed a systematic review of published NMAs assessing biologic therapies for moderate-to-severe psoriasis, aiming to assess the methodological quality of these analyses and explore differences in their results.

What was learned from the study?

Twenty-five NMAs have been published since 2006 and most have come to broadly similar results and conclusions, considering the available data when they were performed.

The most useful NMAs for clinical decision making are those that: include all relevant trials of comparator treatments, assessed in a way that reflects their marketing authorization and expected use in clinical practice; provide a thorough assessment of heterogeneity and inconsistency; and report comparative effects and associated uncertainty in a comprehensive manner.

DIGITAL FEATURES

This article is published with digital features, including a summary slide, to facilitate understanding of the article. To view digital features for this article, go to <https://doi.org/10.6084/m9.figshare.13129880>.

INTRODUCTION

Psoriasis is a chronic, inflammatory, immune-mediated skin disorder [1]. Approximately 70–80% of patients have mild psoriasis, which is adequately managed by topical therapies, whilst the remaining 20–30% present with moderate-to-severe psoriasis [1]. Typically, topical treatments alone can be inadequate for these patients. If there is insufficient improvement in disease symptoms after treatment with phototherapy and/or conventional systemic therapies, patients are offered biologic systemic therapy [1].

The first class of biologic treatments licensed for moderate-to-severe plaque psoriasis were anti-tumor necrosis factor (anti-TNF) agents, such as etanercept, infliximab and adalimumab [2]. Since then, other classes of biologics have been licensed for this indication including a dual interleukin-12/23 (IL-12/23) inhibitor ustekinumab, followed by IL-17 inhibitors (secukinumab, ixekizumab) and receptor antagonist (brodalumab) [3], and IL-23 targeted treatments (guselkumab, tildrakizumab, risankizumab) [4]. Most recently, another anti-TNF treatment, certolizumab pegol, has been licensed in this population [3].

Due to the large number of available treatments for psoriasis, clinical and healthcare funding decisions often require a comparison between available treatment options. Although numerous clinical trials have been undertaken in psoriasis, there is a limited number of head-to-head trials, none of which compare all the available treatment options. In this disease area, similarities in study design and patient populations make these trials particularly suitable for comparison in a network meta-analysis (NMA). This method allows the comparison of multiple interventions in a single coherent analysis so all

available interventions can be compared, using both direct and indirect evidence, thus resulting in a more precise estimate of treatment effect and the ability to rank treatments against one another, even if they have not been compared directly in a head-to-head randomized controlled trial (RCT) [5].

Statistical methods used to carry out NMAs are usually classed as either Bayesian or frequentist, which differ in some of their assumptions, methods for data synthesis, and interpretation of results [6, 7]. Although both analytical frameworks facilitate comparisons of multiple treatments, the Bayesian approach has been used most often, likely due to several advantages over the frequentist [5]. These include the ability to incorporate prior knowledge about the treatment effects into the model, and easy calculation of treatment ranks [8]. It is also the preferred framework of health technology assessors (HTAs) such as the National Institute for Health and Care Excellence (NICE) in the United Kingdom [9].

For both Bayesian and frequentist approaches, two model types can be used when performing an NMA: fixed-effect and random-effects models. A fixed-effect model assumes that all studies are estimating a single effect size and differences observed between study estimates are a result of chance. The results from a fixed-effect model provide an estimate of this underlying effect. The random-effects model assumes different studies are estimating different, but related effects. This is due to the differences between studies in patient and study characteristics (between-study heterogeneity). The results from a random-effects model can be interpreted as a mean of these different effects [5, 10].

Other than the choice of an analytical framework and model type, the results of an NMA may be influenced by a number of factors, such as which trials are included, bias within these trials, inconsistency (differences between direct and indirect estimates of treatment effect), and heterogeneity (and how it is addressed) [11].

As mentioned above, numerous NMAs have been undertaken in psoriasis; however, it is unclear whether these have always used robust

methods or arrived at the same conclusions. Although frequently used to support decision making, there remains skepticism surrounding the use of NMAs, and mistrust of their results [12].

To address these concerns, a systematic literature review (SLR) of the published NMAs assessing biologic treatments for moderate-to-severe plaque psoriasis was performed. The aim was to assess the methodological quality of these analyses and explore the differences in their methods and results. Potential reasons for any differences in results were also considered.

METHODS

A SLR was conducted to identify all NMAs assessing biologic treatments for moderate-to-severe plaque psoriasis. NMAs that evaluated two or more currently licensed biologic treatments in patients with moderate-to-severe psoriasis were included. Analyses evaluating any efficacy or safety outcome were included. Only NMAs using aggregate level data and a Bayesian or frequentist approach were included to allow for coherent comparisons of methodology and results. Indirect treatment comparisons carried out using methods such as those described by Bucher [13], or those using matching-adjustment or simulation methods were excluded because they are more limited in scope. Full eligibility criteria can be found in the supplementary material. The review was carried out in accordance with a protocol developed prior to commencement of work. This article is based on previously conducted studies and does not contain any studies with human participants or animals performed by any of the authors.

Study Identification

Embase, MEDLINE, MEDLINE In-Process, and the Cochrane Library were searched on 13 March 2019 and updated 19 February 2020, without time restrictions. Search strategies included text and index terms for psoriasis, relevant biologic treatments and study design. Full search strategies can be found in the online supplement. Reference lists of included

publications were checked to identify any additional NMAs. After removal of duplicates in EndNote (Thomson Reuters), titles and abstracts were imported into DistillerSR (Evidence Partners) and assessed for inclusion by one reviewer, with a second reviewer independently performing a 40% check. Had the check revealed a significant number of disputes, a full 100% check would have been performed. This did not turn out to be necessary. Full texts were independently assessed for inclusion by two reviewers. Discrepancies were resolved by discussion or, when necessary, by a third reviewer.

Data Extraction and Quality Assessment

Details of each NMA were extracted, including the definition of the question addressed by the NMA, the sources of funding and methods used for identification and selection of studies, data extraction, critical appraisal, and data synthesis. In addition, information was collected on the studies included in each NMA, as well as the base-case results for all treatments compared with placebo, and the ranking of treatments based on efficacy or safety.

Data extraction and subsequent quality assessment of studies was carried out by one reviewer and checked by another. Any disagreements were resolved by discussion or, if necessary, by a third reviewer. Quality assessment was carried out using the International Society of Pharmacoeconomics and Outcomes Research (ISPOR) checklist for assessing reliability of NMAs [14]. This checklist covers five main areas, including (1) the used evidence base, (2) analysis methods, (3) reporting quality and transparency, (4) interpretation of findings, and (5) conflicts of interest [14]. There are 22 questions, allowing readers to better understand the applicability and credibility of the NMA and its results. Similar topic areas are covered in other commonly used assessment checklists, such as one developed by the NICE Decision Support Unit [15].

Data Analysis

The methodological details of identified NMAs were compared in a narrative summary. Psoriasis Area and Severity Index (PASI) response was the most commonly reported efficacy outcome, allowing the broadest comparison between NMAs. For that reason, results for PASI response were the focus of the analysis (safety outcome results were also summarized). A forest plot for each biologic versus placebo was generated using R Studio [16] to present the individual NMA results, with risk ratios (RR) and odds ratios (OR) considered separately. The results of each NMA reporting a treatment effect for the same comparison were compared by visual inspection. Ranking of treatments based on PASI response and safety outcomes was also compared between NMAs, considering the presence of any uncertainty.

RESULTS

Search Results

Electronic searches identified 3043 publications. An additional three were identified through reference checking. After removal of duplicates, a total of 2271 titles and abstracts were assessed for inclusion and 208 papers were assessed in full text: a total of 25 analyses were included. Details can be seen in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram Fig. 1.

Included NMAs

Study Identification and Selection in the NMAs

Table 1 provides brief details of the inclusion criteria in each identified analysis. All analyses included adults with moderate-to-severe psoriasis; Jabbar-Lopez 2017 [17] also included children.

The majority of identified NMAs searched for primary studies in MEDLINE, Embase, and the Cochrane Library. However, in two NMAs only MEDLINE was searched, [18, 19] and in another the search methods were not reported [20]. The

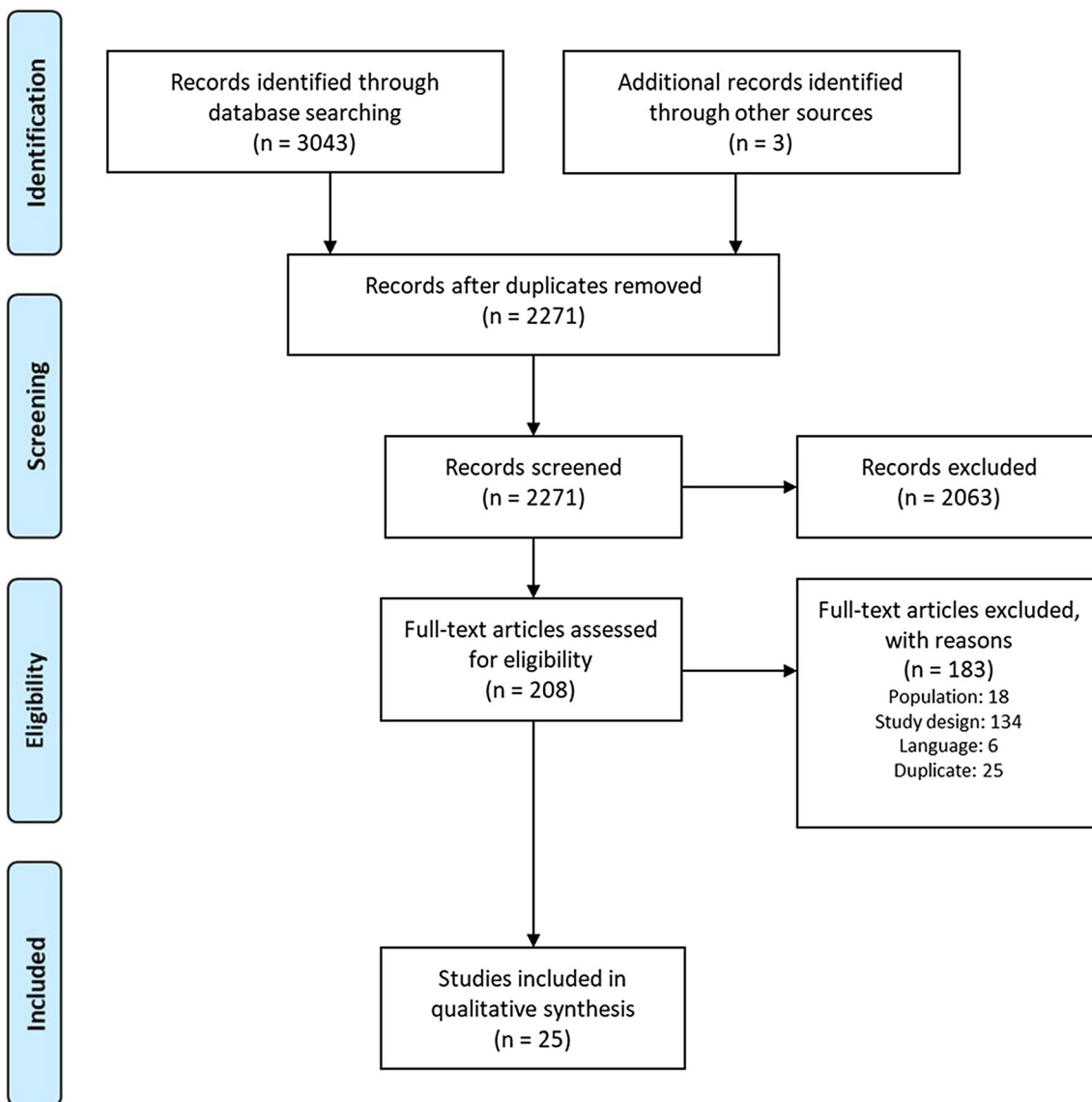


Fig. 1 A PRISMA diagram showing the flow of the included and excluded studies

search strategies in the identified NMAs were generally comprehensive, with the exception of seven analyses, for which the methods used were likely to miss relevant studies [18, 19, 21–26]. Risk of bias was assessed using the Jadad scale for randomized controlled trials [27] in six, the Cochrane Collaboration’s tool for assessing risk of bias [28] in nine, the Newcastle–Ottawa Scale [29] in one and the NICE methodology checklist for RCTs [9] in two NMAs. Six analyses did not clearly report

assessment of risk of bias [20, 22, 23, 30–32]. Details of study identification and selection can be found in the online supplement.

A range of relevant biologic interventions were considered in the analyses, with the total number ranging from four in Geng 2018 [22] to twelve in the Cochrane Reviews by Sbidian and colleagues [33, 34]. The majority of NMAs considered licensed doses. Exceptions included Jabbar-Lopez 2017 [17] and the Cochrane reviews, [33, 34] which included any dose of

Table 1 Study eligibility criteria in the included NMAs

NMA	Interventions		ETN	IFX	ADA	UST	SEC	IXE	BRO	GUS	TIL	RIS	CZP	Other bio	Other non-bio	Population	Outcome	Time point (weeks)	Study design and phase
	L	U3																	
Woolcott 2006 [35]	L	L, U3												✓	✓	Adults, M-to-S PsO	PASI 50, 75, 90	10–12	RCT, phase NR
Reich 2008 [32]	L	L												✓	✓	Active chronic stable plaque PsO	PASI 50, 75, 90	10–12 and 24	RCT, phase 2 and 3
Bansback 2009 [31]	HD	L	L											✓	✓	M-to-S PsO	PASI 50, 75, 90	Not defined; primary endpoint	RCT, phase NR
Reich 2012 [39]	LD	L	L	L	WB, U45	U90								✓	✓	Adults, plaque type PsO	PASI 50, 75, 90	10–16	RCT, phase NR
Lin 2012 [43]	HD	L	L	L	U45, U90									✓	✓	M-to-S plaque PsO	PASI 50, 75, 90	10–16	RCT, phase 3
Gupta 2014 [19]	L	L	L	L	U45									✓	✓	Adults (18+), chronic plaque PsO	PASI 75	10–16	RCT, phase 3 and above
Messori 2015 [18]	L	L	L	L	U45, U90											M-to-S PsO	Safety: AEs, Infectious AEs	12–24	RCT, phase NR
Signorovitch 2015 [42]	L	L	L	L	U45, U90									✓	✓	M-to-S PsO	PASI 50, 75, 90	10–16	RCT, phase 2 and 3
Gomez Garcia 2016 [23]	L	L	L	L	WB, U45	U90	L									Adults, M-to-S PsO	PASI 75, 90; Safety: any AE	10–16	RCT, phase NR
Jabbar-Lopez 2017 [17]	A	A	A	A	A	A	A	A	A	A	A	A	A			Adults and children, any severity PsO; ≤ 50% PsA	PASI 75; withdrawal due to AE; DLQI; Clear/almost clear	> 12 (but include 10-week IFX)	RCT, phase 2 and 3; ≥ 50 patients
Sbidian 2017 [34]	A	A	A	A	A	A	A	A	A	A	A	A	A	✓	✓	Adults, M-to-S plaque PsO OR PsA with M-to-S PsO	PASI 75, 90; SAE, AE; PGA 0/1; DLQI	12–16	RCT, phase 2 and above
Geng 2018 [22]	L, U25	L, U3	L	L	U45, U90									✓	✓	Adults, M-to-S PsO	PASI 50, 75, 90	NR	RCT, phase NR

Table 1 continued

NMA	Interventions										Population	Outcome	Time point (weeks)	Study design and phase		
	ETN	IFX	ADA	UST	SEC	IXE	BRO	GUS	TIL	RIS					CZP	Other bio
Loos 2018 [24]	HD	L	L	L	WB, U45, U90	L	L	L	L	L	L	✓	Adults, M-to-S chronic plaque PsO	PASI 50, 75, 90	10–16	RCT, phase 3
Sawyer 2018 (i) [36]	LD, HD	L	L	L	WB, U45, U90	L	L	L	L	U140	L	✓	Adults, M-to-S chronic plaque PsO	PASI 50, 75, 90, 100	10–16	RCT, phase NR
Sawyer 2018 (m) [40]	HD	L	L	L	WB, U45, U90	L	L	L	L	L	L	✓	Adults, M-to-S chronic plaque PsO	PASI 75, 90, 100	40–64	RCT, phase 2 and 3, OLE
Cameron 2018 [38]	L	L	L	L	WB, U45, U90	L	L	L	L	L	L ^a	✓	Adults, M-to-S chronic plaque PsO	PASI 50, 75, 90, 100; PGA 0/1; safety	End of induction	RCT, phase 3 and 4
Sawyer 2019 [37]	LD, HD	L	L	L	WB, U45, U90	L	L	L	L	L100	L	✓	Adults, M-to-S chronic plaque PsO	PASI 50, 75, 90, 100	10–16	RCT, phase NR
Bai 2019 [21]					U45, WB + U90	L	L	L	L	L	L	✓	Adults, M-to-S chronic plaque PsO	PASI 75, 100; PGA 0/1; AE; SAE; discontinuation due to AE	12–16	RCT, phase NR
Xu 2019a ^b [25]	HD	L	L	L	A	A	A	L	L	L	L	✓	Adults, M-to-S chronic plaque PsO	PASI 50, 75, 90, 100; PGA 0/1; DLQI; headache; infection; discontinuation;	12–16	RCT or quasi RCT, phase NR
Xu 2019b [26]			L	WB	A	A	A	A	A	A	A		Adults, M-to-S chronic plaque PsO	PASI 75; AE	12–24	RCT, phase NR
Warren 2019 [41]	LD, HD	L	L	L	WB, U45, U90	L	L	L	L	L100, L200	L		M-to-S chronic plaque PsO	PAIS 75, 90, 100	12–16	RCT, phase 2 and above
Wu 2020 [44]	A	A	A	A	A	A	A	A	A	A	A ^a	✓	Patients with PsO	Change in body weight or BMI	Up to 24	RCT or non-RCT

Table 1 continued

NMA	Interventions													Population	Outcome	Time point (weeks)	Study design and phase
	ETN	IFX	ADA	UST	SEC	IXE	BRO	GUS	TIL	RIS	CZP	Other bio	Other non-bio				
Sbidian 2020 [33]	A	A	A	A	A	A	A	A	A	A	A	✓	✓	Adults, M-to-S chronic plaque PsO	PASI 75, 90; SAE, AE; PGA 0/1; DLQI	8–24	RCT, phase 2 and above
Warren 2020 [20]	HD	L	L	WB	L	L	L	L	L100	L	L400			Adults, M-to-S chronic plaque PsO	PASI 75, 90, 100; DLQI 0/1;	4, 8 and 12	RCT, phase 3
Armstrong 2020 [30]	L	L	L	WB, U45, U90	L	L	L	L	L100, L200	L	L200, L400	✓	✓	Adults, M-to-S chronic plaque PsO	PASI 75, 90, 100	10–16; 44–60	RCT, phase 2 and above

ADA adalimumab, *AE* adverse event, *BMI* body mass index, *BRO* brodalumab, *CZP* certolizumab pegol, *DLQI* dermatology life quality index, *ETN* etanercept, *GUS* guselkumab, *i* induction, *IFX* infliximab, *IXE* ixekizumab, *m* maintenance, *M-to-S* moderate- to -severe, *NR* not reported, *OLE* open-label extension, *PASI* psoriasis area severity index, *P&A* psoriatic arthritis, *P&A* physicians global assessment, *PsO* psoriasis, *RCT* randomised control trial, *RIS* risankizumab, *SAE* serious adverse event, *SEC* secukinumab, *TIL* tildrakizumab, *UST* ustekinumab, *A* any dose, *HD* high etanercept dose of 100 mg/week, *kg* kilogram, *L* Licensed dose(s), *LD* low etanercept dose of 50 mg/week, *L100* licensed dose of 100 mg, *L200* licensed dose of 200 mg, *L400* licensed dose of 400 mg, *mg* milligram, *U* unlicensed dose(s), *U3* unlicensed infliximab dose of 3 mg/kg, *U25* unlicensed etanercept dose of 25 mg/week, *U45* ustekinumab at 45 mg irrespective of patient's body weight, *U90* ustekinumab at 90 mg irrespective of patient's body weight, *U140* unlicensed brodalumab dose of 140 mg, *WB* ustekinumab at 45 mg for patients with body weight up to 100 kg and 90 mg for patients with body weight of 100 kg or more

^a Included in protocol, no evidence was identified; other non-biologic includes apremilast

^b The labels in this study suggest that only licensed doses were included; however, closer inspection reveals that some study data relates to trial arms of unlicensed doses. It is assumed that these were combined with studies evaluating licensed doses

treatments of interest; Geng 2018, [22] which included unlicensed doses of etanercept and infliximab; Woolacott 2006, [35] which included unlicensed doses of infliximab; and Sawyer 2018 (induction [i]) [36], Sawyer 2019, [37] and Cameron 2018 [38] which included unlicensed doses of several therapies where their inclusion added indirect evidence. Two unlicensed doses of ustekinumab were also commonly included in analyses: 45 mg and 90 mg, irrespective of patient's body weight. Eleven of 22 NMAs evaluating ustekinumab included the licensed weight-based dose [20, 23, 24, 26, 30, 36–41].

The most frequently assessed outcomes were PASI response in 23 NMAs, followed by safety in nine NMAs. Seven analyses assessed other efficacy or quality of life outcomes [17, 20, 21, 25, 33, 34, 38]. All but one of the 25 analyses evaluated treatments at the end of the induction phase (between 10 and 24 weeks, depending on the analysis), while Sawyer 2018 (maintenance [m]) and Armstrong 2020 (m) compared treatments at one year. All NMAs included RCTs; Sawyer 2018 (m) and Armstrong 2020 (m) additionally included long-term extensions of RCTs and Wu 2020 also included non-randomized studies. Eight NMAs reported including phase 2 studies and above [17, 30, 32–34, 40–42], five NMAs restricted inclusion to just phase 3 studies and above [19, 20, 24, 38, 43] and the rest did not specify.

Source of Funding and Conflicts of Interest

Only one analysis [22] did not report on conflicts of interest or source of funding. Two NMAs declared no conflicts of interest including any sources of funding [18, 25], two declared conflicts but no source of funding, [43, 44] and two reported on potential conflicts but not on sources of funding [19, 21]. The remaining eighteen provided details of conflicts of interests and the funding received. Seven NMAs [17, 23, 24, 26, 33–35] were funded by public grants (such as from the Institute for Clinical and Economic Review [24] or the Cochrane Collaboration [33, 34]) and eleven by drug manufacturers [20, 30–32, 36–42].

NMA Analytical Methods and Assessment of Quality using ISPOR Criteria

The methods used in each of the NMAs can be seen in Table 2. Nineteen NMAs used a Bayesian approach and six used a frequentist approach. PASI response was assessed in 23 NMAs; Messori 2015 [18] assessed safety outcomes only and Wu 2020 [44] assessed the impact of biologic therapy on body weight and body mass index. Ten NMAs evaluated PASI as an ordered categorical outcome, eleven treated PASI response as a dichotomous outcome, and one analyzed the outcome both ways [38]. One analysis compared biologic therapies on their cumulative clinical benefit, measured by the area under the curve (AUC) for PASI 75, 90, and 100 [41].

Most analyses considered each dose of a treatment separately, or only combined different dosing schedules that resulted in the same weekly dose (for example, etanercept 25 mg twice a week and 50 mg once a week). However, Jabbar-Lopez 2017 [17] and the Cochrane Reviews [33, 34] included any evaluated dose (licensed or unlicensed, including phase 2 doses not tested in phase 3) and combined (pooled) these in the analysis.

The details of critical appraisal of the NMAs using the ISPOR checklist can be seen in Table 3. Seven analyses were judged to have performed systematic reviews that could have failed to identify and include all relevant RCTs, [18, 19, 21, 22, 24–26] and one study did not report details on how included studies were identified and included [20]. The quality of primary studies included in the NMAs was generally reported as good, minimizing the risk of bias, with the exception of non-RCT studies included by Wu et al. [44] There appeared to be no evidence of selective reporting of results in primary studies included in the identified NMAs.

In all NMAs the identified studies formed a connected network. All analyses used statistical methods that preserve within-study randomization, with the exception of Armstrong 2020 (m), which used an unanchored indirect comparison.

The rationale for the choice of a fixed-effect or random-effects model was discussed in nine

Table 2 Methods employed by each included network meta-analysis

NMA	Analysis and base-case model type	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Woolacott 2006 [35]	Bayesian, ordered probit, RE	Ordered categorical (50, 75, 90); RR	4–16 weeks	ETN doses analyzed separately, IFX combined 5 mg and 3 mg/kg ^a	Qualitative assessment of clinical characteristics and quantitative assessment based on χ^2 for pair-wise analyses; RE model used	N/A; no loops of evidence in the network	16
Reich 2008 [32]	Bayesian, ordered probit, RE	Ordered categorical (50, 75, 90); RR	10–12 and 24 weeks	Analyzed separately	Assessed based on clinical characteristics and χ^2 for pair-wise analysis; only considered significant for infliximab; explored by removing the smallest study from pairwise analyses; RE model used	N/A; no loops of evidence in the network	15
Bansback 2009 [31]	Bayesian, ordered probit, RE	Ordered categorical (50, 75, 90); RR	8–16 weeks	One dose per drug with lower or higher doses tested in sensitivity analysis	Qualitative assessment of baseline characteristics; meta-regression on baseline characteristics considered, but deemed unnecessary; sensitivity analyses performed using different trials and high or low evaluated dosages were run and compared with base case results; RE model used	NR	22

Table 2 continued

NMA	Analysis and base-case model type	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Reich 2012 [39]	Bayesian, ordered probit, RE	Ordered categorical (50, 75, 90); RR	10–16 weeks	Analyzed separately	Assessment NR; RE model used	NR	20
Lin 2012 [43]	Bayesian, binomial logit (assumed), RE	Dichotomous (50, 75, 90); OR	10–16 weeks	Analyzed separately	Three models assessed for best fit: FE, RE, and RE with meta-regression on disease duration and BSA involvement; sensitivity analyses performed using high or low evaluated dosages and excluding studies with only treatment-naïve patients; RE model used	Compared results including and excluding head-to-head trials, but results inconclusive	17
Gupta 2014 [19]	Bayesian, binomial logit (assumed), RE	Dichotomous (75); OR	10–16 weeks	For ETN both licensed doses combined; for other one dose per drug	Qualitative assessment of study and patient characteristics; quantitative assessment based on Q and I ² statistics for pairwise analyses; RE model used	Node-splitting analysis for each loop of evidence in network; no inconsistency detected	21

Table 2 continued

NMA	Analysis and base-case model type	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Messori 2015 [18]	Bayesian; binomial logit (assumed), FE	N/A	12–24 weeks	Analyzed separately	Accounted for in a meta-regression (not described); RE and FE models compared for goodness of fit and FE model used	NR	Any SAE 10, any infectious AE 11
Signorovitch 2015 [42]	Bayesian, ordered probit, RE, included regression adjustment for placebo arm responses	Ordered categorical (50, 75, 90); RR	8–16 weeks	Analyzed separately	RE models with and without adjustment for placebo arm responses compared for goodness of fit and RE model with adjustment used; sensitivity analyses performed using WB dosing for ustekinumab and using binomial logit model at PASI 75 alone	NR	15
Gomez Garcia 2016 [23]	Frequentist multivariate, RE	Dichotomous (75, 90); OR	10–16 weeks	Analyzed separately	Quantitative assessment based on I^2 in pairwise meta-analysis	Assessed using the ratio of ORs where a value ≥ 2 indicated potential inconsistency; some evidence of inconsistency identified	27

Table 2 continued

NMA	Analysis and base-case model type	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Jabbar-Lopez 2017 [17]	Frequentist multivariate, RE	Dichotomous (75, 90 ^b); OR	> 12 weeks reported, but 10-week IFX trials included	All doses combined	Assessed by visual inspection of forest plots; RE model used; subgroup analysis included data on licensed biologic doses only	Assessed by visual inspection of forest plots, loop-specific inconsistency plots and calculation of an inconsistency factor ^b ; some evidence of inconsistency identified	PASI 75: 40; Withdrawals due to AEs: 36
Sbidian 2017 [34]	Frequentist multivariate, choice of FE or RE NR	Dichotomous (75, 90); RR	12–16 weeks	All doses combined	Assessed by inspection of patient and study characteristics and estimation of between-study standard deviation followed by comparison with empirical distributions; subgroup analyses and meta-regression on key patient characteristics planned, but data insufficient to perform; sensitivity analyses at dose-level considering each drug dose a different intervention, excluding trials at high risk of bias or with total sample size of < 50 patients and analyzing outcomes on observed cases only	Assessed using loop-specific and side-splitting approaches (Bucher 1997 and Dias 2010) and by fitting a design by treatment interaction model; some evidence of inconsistency identified	PASI 90: 58; SAE: 60; PASI 75: 64; Any AE: 54

Table 2 continued

NMA	Analysis and base-case model type	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Geng 2018 [22]	Bayesian, binomial logit (assumed), RE	Dichotomous assumed (50, 75, 90); OR	NR	Analyzed separately	Assessment NR; RE model used	Assessed by node-splitting; because $p > 0.05$, the inconsistency model was used	33
Loos 2018 [24]	Bayesian, ordered probit, RE, included regression adjustment for placebo arm responses	Ordered categorical (50, 75, 90); RR	10–16 weeks	Only one dose per drug except UST, where doses were combined	Assessment NR; RE model with adjustment for placebo arm responses used	NR	34
Sawyer 2018 (i) [36]	Bayesian, ordered probit, RE, included regression adjustment for placebo arm responses	Ordered categorical (50, 75, 90, 100); RR	10–16 weeks	Analyzed separately	Qualitative assessment of study and patient characteristics; RE and FE models with and without adjustment for placebo arm responses compared for goodness of fit and RE model with adjustment used; sensitivity analyses performed including data from trial arms of unlicensed doses of systemic therapies	Assessed using RE unrelated mean effects model; no evidence of inconsistency detected	41

Table 2 continued

NMA	Analysis and base-case model type	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Sawyer 2018(m) [40]	Bayesian, ordered probit, FE	Ordered categorical (75, 90, 100); RR	40–64 weeks	Analyzed separately	Qualitative assessment of study design, patient characteristics and endpoints; FE and RE models assessed for goodness of fit and FE model used; sensitivity analysis performed excluding studies with < 5% of patients report prior biologic exposure	Assessed using loop-specific approach (Bucher 1997); no evidence of inconsistency detected	17

Table 2 continued

NMA	Analysis and base-case model type	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Cameron 2018 ^c [38]	Base-case analysis Bayesian, binomial logit, RE, included regression adjustment for placebo arm responses	Dichotomous (50, 75, 90, 100); RR	10–16 weeks	Analyzed separately	Qualitative assessment of study and patient characteristics; quantitative assessment based I ² statistics for pairwise analyses; FE and RE models combined with several meta-regressions to account for variation in placebo response, prior biologic use, body weight, disease duration, age, race, and baseline outcome scores were compared for goodness of fit; RE model with placebo adjustment used for PASI 50, 90 and 100 and RE model with disease duration adjustment used for PASI 75	Assessed by comparing the mean of the posterior residual deviance and DIC statistics in fitted consistency and inconsistency models; no evidence of inconsistency detected	45
	Supplementary analysis ^c Bayesian, ordered probit, RE, included regression adjustment for placebo arm responses	Ordered categorical (50, 75, 90, 100); RR			RE model combined with meta-regression to account for variation in placebo response using covariate at each PASI cut-off; RE model with placebo adjustment used		

Table 2 continued

NMA type	Analysis and base-case model	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Sawyer 2019 [37]	Bayesian, ordered probit, RE, included regression adjustment for placebo arm responses	Ordered categorical (50, 75, 90, 100); RR	10–16 weeks	Analyzed separately	Qualitative assessment of study and patient characteristics; RE, models with and without adjustment for placebo arm responses compared for goodness of fit; sensitivity analyses performed to test impact of excluding trials reporting < 5% biologic exposed patients, using a different timepoint (12 weeks rather than 16 weeks) for secukinumab, and excluding studies with fewer than 50 patients per arm	Assessed using RE unrelated mean effects model; no evidence of inconsistency detected	77
Bai 2019 [21]	Frequentist, model type NR, choice of FE or RE NR	Dichotomous (75, 100); RR	12–16 weeks	Only one dose per drug except UST, where high dose and WB dose combined	NR	Assessed using loop-specific approach; some inconsistency detected	28

Table 2 continued

NMA	Analysis and base-case model type	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Xu 2019a [25]	Bayesian, binomial logit (assumed), RE	Dichotomous (50, 75, 90, 100); OR	12–16 weeks	Only one dose per drug except IXE, SEC, UST, where doses were combined ^d	Assessment NR; RE model used	Node-splitting and net heat plots conducted; node-splitting showed some inconsistency; net heat plots showed no statistical inconsistency	54
Xu 2019b [26]	Bayesian, binomial logit (assumed), choice of FE or RE NR	Dichotomous (75); RR	12–24 weeks	All doses combined	Assessment NR; subgroup analyses stratified by drugs and time of follow-up. Meta-regression analysis was performed by publication year, drugs, number of patients, male ratio, white ratio, time of follow-up, biologic agents and PsA. Sensitivity analysis was conducted to find whether one or more studies deviated from the overall results	Method NR, but no significant inconsistency found	13

Table 2 continued

NMA type	Analysis and base-case model	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Warren 2019 [41]	Bayesian, model type NR, FE	Unclear; proportion of AUC (75, 90, 100)	12–16 weeks	Analyzed separately	Authors comment on similarity of inclusion/exclusion criteria of trials and similarity of baseline characteristics; FE and RE models assessed for goodness of fit and FE model used; estimated percentages of maximum AUC from NMA were placebo-adjusted, but no rationale provided for why	NR	28
Wu 2020 [44]	Frequentist, model type NR, RE	N/A	Up to 24 weeks	All doses combined	NR	Assessed using design-by-treatment model, loop inconsistency model and node-splitting model	BW change: 6 BMI change: 5

Table 2 continued

NMA	Analysis and base-case model type	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Sbidian 2020 [34]	Frequentist multivariate, choice of FE or RE NR	Dichotomous (75, 90); RR	8–24 weeks	All doses combined	Assessed by inspection of patient and study characteristics and estimation of between-study standard deviation followed by comparison with empirical distributions; subgroup analyses and meta-regression on key patient characteristics planned, but data insufficient to perform; sensitivity analyses at dose-level considering each drug dose a different intervention, excluding trials at high risk of bias or with total sample size of < 50 patients and analyzing outcomes on observed cases only	Assessed using loop-specific and side-splitting approaches (Bucher 1997 and Dias 2010) and by fitting a design by treatment interaction model; some evidence of inconsistency identified	PASI 75: 107 PASI 90: 95
Warren 2020 [20]	Bayesian and frequentist, model type NR, FE	Dichotomous (75, 90); unclear	4, 8 and 12 weeks	Analyzed separately	Assessment NR; authors assumed studies to be generally similar and consistent	NR	33

Table 2 continued

NMA	Analysis and base-case model type	PASI outcome type (cut-offs); summary statistic	Timepoints	Treatment of doses	How was heterogeneity considered	How was inconsistency considered	Number of studies
Armstrong 2020 [30]	Bayesian, ordered probit, RE, included regression adjustment for placebo arm responses	Ordered categorical (50, 75, 90, 100); Response rates	10–16 weeks	Analyzed separately	Assessment NR; to account for heterogeneity across trials, a model that adjusted for reference-arm response was implemented	NR	60

ADA adalimumab, *AE* adverse events, *AUC* area under the curve, *BMI* body mass index, *BSA* body surface area, *BW* body weight, *ETN* etanercept, *FE* fixed-effect (model), *IFX* infliximab, *IXE* ixekizumab, *kg* kilogram, *i* induction, *m* maintenance, *mg* milligram, *N/A* not applicable, *NMA* network meta-analysis, *NR* not reported, *OR* odds ratio, *PASI* Psoriasis Area and Severity Index, *PsA* psoriatic arthritis, *RE* random-effects (model), *RR* risk ratio, *SAE* serious adverse events, *SEC* secukinumab, *UST* ustekinumab, *WB* weight-based

^a The same outcome code was used in the model

^b Inconsistency factor was defined as the logarithm of the ratio of two ORs from direct and indirect evidence in a loop, with values close to 1 suggesting agreement between direct and indirect evidence

^c The only results available in the base-case analysis were for guselkumab versus each comparator; comparisons for all treatments versus placebo were available in a further analysis, supplied in the online supplement. These results are used for the comparisons in our analysis

^d The labels in this study suggest that only licensed doses were included; however, closer inspection reveals that the study data relates to trial arms of unlicensed doses. It is assumed that these were combined with studies evaluating licensed doses

Table 3 Assessment of the quality of included network meta-analyses using the ISPOR checklist

Question	Woolacott 2006 [35]	Reich 2008 [32]	Bansback 2009 [31]	Reich 2012 [39]	Lin 2012 [43]	Gupta 2014 [19]	Messori 2015 [18]	Signorovitch 2015 [42]	Gomez Garcia 2016 [23]	Jabbar- Lopez 2017 [17]	Sbidian 2017 [34]	Geng 2018 [22]	Loos 2018 [24]
Did the researchers attempt to identify and include all relevant randomized controlled trials? (RCTs)	✓	✓	✓	✓	✓	✗	✗	✓	✗	✓	✓	✗	✗
Do trials for interventions of interest form one connected network of RCTs?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Is it apparent that poor quality studies were included, thereby leading to bias?	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗
Is it likely that bias was induced by selective reporting of outcomes in the studies?	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗

Table 3 continued

Question	Woolacott 2006 [35]	Reich 2008 [32]	Bansback 2009 [31]	Reich 2012 [39]	Lin 2012 [43]	Gupta 2014 [19]	Messori 2015 [18]	Signorovitch 2015 [42]	Gomez Garcia 2016 [23]	Jabbar-Lopez 2017 [17]	Sbidian 2017 [34]	Geng 2018 [22]	Loos 2018 [24]
Are there systematic differences in treatment effect modifiers across the different treatment comparisons in the network?	X	X	X	X	X	X	?	✓	✓	✓	✓	✓	✓
Were imbalances in effect modifiers across the different treatment comparisons identified before reviewing individual study results?	X	X	X	X	✓	✓	X	✓	X	X	✓	X	✓
Were statistical methods used that preserve within-study randomization?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 3 continued

Question	Woolacott 2006 [35]	Reich 2008 [32]	Bansback 2009 [31]	Reich 2012 [39]	Lin 2012 [43]	Gupta 2014 [19]	Messori 2015 [18]	Signorovitch 2015 [42]	Gomez 2016 [23]	Jabbar- Lopez 2017 [17]	Sbidian 2017 [34]	Geng 2018 [22]	Loos 2018 [24]
If both direct and indirect comparisons are available for pairwise contrasts was agreement evaluated or discussed?	N/A	N/A	X	X	✓	✓	X	X	✓	✓	✓	✓	X
In the presence of consistency, were both direct and indirect evidence included in the network meta-analysis?	NI	✓	NI	NI	✓	✓	NI	NI	✓	✓	✓	✓	NI
With an imbalance in the distribution of treatment effect modifiers across the different types of comparisons in the network of trials, did researchers attempt to minimize this bias with the analysis?	N/A	N/A	N/A	N/A	✓	N/A	NI	✓	NI	NI	✓	NI	✓

Table 3 continued

Question	Woolacott 2006 [35]	Reich 2008 [32]	Bansback 2009 [31]	Reich 2012 [39]	Lin 2012 [43]	Gupta 2014 [19]	Messori 2015 [18]	Signorovitch 2015 [42]	Gomez 2016 [23]	Jabbar- Lopez 2017 [17]	Sbidian 2017 [34]	Geng 2018 [22]	Loos 2018 [24]
Was a valid rationale provided for the use of RE or FE models?	✓	✓	✗	✗	✓	✗	✓	✗	✗	✗	✓	✗	✓
If a RE model was used, were assumptions about heterogeneity explored or discussed?	✗	✗	✗	✗	✓	✓	N/A	✓	✗	✗	✓	✗	✓
Is a graphical or tabular representation of the evidence network provided with information on the number of RCTs per direct comparison?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Are the individual study results reported?	✓	✓	✓	✓	P	✗	✓	✓	✗	✓	✓	✓	✓

Table 3 continued

Question	Woolacott 2006 [35]	Reich 2008 [32]	Bansback 2009 [31]	Reich 2012 [39]	Lin 2012 [43]	Gupta 2014 [19]	Messori 2015 [18]	Signorovitch 2015 [42]	Gomez 2016 [23]	Jabbar- Lopez 2017 [17]	Sbidian 2017 [34]	Geng 2018 [22]	Loos 2018 [24]
Are direct results reported separately from results of indirect comparisons or network meta-analyses?	P	✓	X	X	P	P	X	X	✓	✓	✓	P	X
Are all pairwise contrasts between interventions as obtained with NMA reported along with measure of uncertainty?	X	X	X	X	P	✓	✓	P	✓	✓	✓	✓	✓
Is a ranking of interventions provided given the reported treatment effects and its uncertainty outcome?	X	X	X	✓	P*	P*	✓	X	✓	✓	✓	P*	X
Is the impact of important patient characteristics on treatment effects reported?	X	X	X	✓	✓	X	X	✓	X	X	✓	X	X

Table 3 continued

Question	Sawyer 2018 (i) [36]	Sawyer 2018 (m) [40]	Cameron 2018 [38]	Sawyer 2019 [37]	Bai 2019 [21]	Xu 2019a [25]	Xu 2019b [26]	Warren 2019 [41]	Wu 2020 [44]	Sbidian 2020 [33]	Warren 2020 [20]	Armstrong 2020 [30]
Did the researchers attempt to identify and include all relevant randomized controlled trials (RCTs)?	✓	✓	✓	✓	✗	✗	✗	✓	✓	✓	?	✓
Do trials for interventions of interest form one connected network of RCTs?	✓	✓	✓	✓	✓	✓	✓	✓	P	✓	✓	✓
Is it apparent that poor quality studies were included, thereby leading to bias?	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗
Is it likely that bias was induced by selective reporting of outcomes in the studies?	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
Are there systematic differences in treatment effect modifiers across the different treatment comparisons in the network?	✓	✓	✓	✓	✗	✓	✗	✓	?	✓	?	✓
Were imbalances in effect modifiers across the different treatment comparisons identified before reviewing individual study results?	✓	✓	✓	✓	✗	✗	✗	✗	✗	✓	✗	✓
Were statistical methods used that preserve within-study randomization?	✓	✓	✓	✓	✓	✓	✓	✓	N/A	✓	✓	✓

Table 3 continued

Question	Sawyer 2018 (i) [36]	Sawyer 2018 (m) [40]	Cameron 2018 [38]	Sawyer 2019 [37]	Bai 2019 [21]	Xu 2019a [25]	Xu 2019b [26]	Warren 2019 [41]	Wu 2020 [44]	Sbidian 2020 [33]	Warren 2020 [20]	Armstrong 2020 [30]
If both direct and indirect comparisons are available for pairwise contrasts was agreement evaluated or discussed?	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗	✗
In the presence of consistency, were both direct and indirect evidence included in the network meta-analysis?	✓	✓	✓	✓	✓	✓	✓	NI	✓	✓	NI	NI
With an imbalance in the distribution of treatment effect modifiers across the different types of comparisons in the network of trials, did researchers attempt to minimize this bias with the analysis?	✓	✓	✓	✓	NI	NI	NI	NI	NI	✓	NI	✓
Was a valid rationale provided for the use of RE or FE models?	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗
If a RE model was used, were assumptions about heterogeneity explored or discussed?	✓	✓	✓	✓	✗	✗	NR	N/A	✗	✓	NR	✓

Table 3 continued

Question	Sawyer 2018 (i) [36]	Sawyer 2018 (m) [40]	Cameron 2018 [38]	Sawyer 2019 [37]	Bai 2019 [21]	Xu 2019a [25]	Xu 2019b [26]	Warren 2019 [41]	Wu 2020 [44]	Sbidian 2020 [33]	Warren 2020 [20]	Armstrong 2020 [30]
Is a graphical or tabular representation of the evidence network provided with information on the number of RCTs per direct comparison?	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓
Are the individual study results reported?	✓	✓	✗	✓	✓	✗	✓	✓	✓	✓	✗	✓
Are direct results reported separately from results of indirect comparisons or network meta-analyses?	✗	✗	✗	✗	✗	✗	✓	✗	✓	✓	✗	✗
Are all pairwise contrasts between interventions as obtained with NMA reported along with measure of uncertainty?	✗	✓	✓	✓	✓	✓	P	✗	✓	✓	✗	✗
Is a ranking of interventions provided given the reported treatment effects and its uncertainty outcome?	✗	✓	✓	✗	✓	✓	✓	✗	✓	✓	P*	✗
Is the impact of important patient characteristics on treatment effects reported?	✓	✓	✓	✓	✗	✗	✓	✗	✗	✓	✗	✓

N/A not applicable, NI not investigated, NR not reported, P partly, ? unclear, * no uncertainty reported, ✗ no; ✓ yes

analyses [18, 24, 32, 34–36, 38, 40, 43]. Of the seventeen NMAs using a random-effects model, assumptions about heterogeneity were explored in eleven. Imbalances in effect modifiers across treatment comparisons were considered in 11 analyses; methods that aimed to reduce the impact of bias due to these imbalances were used in ten analyses [24, 30, 33, 34, 36–38, 40, 42, 43]. Seven analyses considered the impact of patient characteristics on treatment effects in a qualitative fashion [34, 36, 38–40, 42, 43].

All analyses included both direct and indirect evidence in their evidence networks. The consistency of direct and indirect evidence was discussed in only fifteen (details of methods used are provided in Table 2). Two analyses [32, 35], however, did not need to assess inconsistency as there were no closed loops in their network.

Reporting of NMA results was not always complete. All but five NMAs presented the results of individual primary studies, but the majority did not present the direct comparison results separately from the indirect comparison or NMA results. Fourteen analyses reported results of all pairwise comparisons from the NMA for the outcome of interest and 16 reported treatment ranking in some form. Three studies [19, 22, 43] reported only the rank order of evaluated therapies based on comparative effect sizes and one reported the mean rank, but not its associated uncertainty [20]. Three studies presented rank results as “rankogram” plots, illustrating the probability distribution for each treatment’s rank [18, 25, 39]. Nine studies [17, 21, 23, 26, 33, 34, 38, 40, 44] presented overall rankings based on the Surface Under the Cumulative RAnking curve (SUCRA), which is a numeric representation of overall rank [45]. Higher SUCRA values (bounded at 100%) indicate that a therapy is ranked best or near the best, and lower values (bounded at 0%) indicate that a therapy is ranked among the worst.

Comparison of Primary Studies Included in NMAs

The trials included in each NMA can be seen in the supplementary material. Overall, the studies included in each NMA were similar, considering the interventions of interest and the search dates. The studies included in each NMA largely reflected the objective of the analysis, quality of the search, and subsequent choices made at the protocol stage. Some NMAs included evidence for oral systemic therapies like apremilast, fumaric acid esters, methotrexate and cyclosporine; therefore, placebo and head-to-head evidence for these therapies were included. Many NMAs excluded phase 2 studies or imposed a minimum trial sample size. Jabbar-Lopez 2017 [17] was the only study to include studies in pediatric patients and to exclude studies with fewer than 50 participants. The 2020 Cochrane Review included studies focused on nail psoriasis [46, 47], whereas these studies did not meet the inclusion criteria for other NMAs. Similarly, two RCTs [48, 49] focusing on patients with psoriatic arthritis appear to have been included in a pair of NMAs [22, 25] but nowhere else.

Another discrepancy between NMAs was the inclusion of studies evaluating only unlicensed doses, such as phase 2 dose-ranging studies of ixekizumab [50], secukinumab [51, 52], risankizumab [53] and guselkumab, [54, 55] which were included in the Cochrane Reviews [33, 34], Jabbar-Lopez, [17] and three others [21, 25, 26].

Further differences can be seen for infliximab trials, many of which were excluded in the 2017 Cochrane Review [34] due to their assessment of outcomes at 10 weeks. In their 2020 update, the timepoint criterion was relaxed and studies lasting at least 8 weeks were included [33]. Jabbar-Lopez [17] stated the exclusion of 10-week trials, but in fact included infliximab data coming from these.

NMA Results

PASI Response

Fourteen analyses reported results for PASI 90; 19 for PASI 75; 11 for PASI 50 and 5 for PASI



Fig. 2 Risk Ratios versus placebo at PASI 90 across NMAs

100. Comparative effects were presented as RRs, ORs, numbers needed to treat (NNT) and proportion of maximum AUC [41]. Results of NMAs for PASI 90 RR can be found in Fig. 2. The results

for OR and other PASI levels are included in the supplementary material.

All analyses reported the biologic treatments to be significantly more efficacious than placebo. Across all analyses using RRs, efficacy

estimates were broadly similar for each intervention. The major exceptions were the 2017 and 2020 Cochrane Reviews [33, 34], which provided lower estimates of efficacy than other NMAs for each level of PASI and for each intervention. The results measured using an OR appeared to be mostly consistent. One major difference was seen for PASI 90 response to ustekinumab compared with placebo, where Jabbar-Lopez 2017 [17] provided a substantially lower OR than the other NMAs; and for the infliximab compared with placebo, where Xu 2019a reported a very high upper limit of the 95% confidence interval [25].

For older biologics, in particular infliximab and etanercept, a trend was observed whereby efficacy estimates reported by each analysis increased over time until the analysis by Signorovitch 2015 [42], when they begin to decrease; a plot depicting this can be found in the supplementary material.

Table 4 shows the treatment ranking across the included analyses. Ranking was reported in 14 NMAs. Ten analyses reported ranking on PASI 75, six on PASI 90, and one on all levels of PASI. For the remaining analyses and where rank across multiple categories was reported, ranking was inferred based on the results for PASI 90, if available, and on PASI 75 if not. Two analyses did not report PASI response and are therefore not included in this Table [18, 44].

The treatment ranking appears consistent across analyses published at a similar time. As newer biologics became available, these were generally ranked higher than older therapies except infliximab which continues to rank amongst the highly efficacious treatments, when assessing short-term efficacy. In recent NMAs the most efficacious treatments, at PASI 90, were the IL-17 inhibitors (brodalumab, ixekizumab, secukinumab), IL-23 inhibitors (guselkumab and risankizumab), and infliximab. All of these treatments tend to be significantly more efficacious than adalimumab, certolizumab, etanercept, and ustekinumab. As far as differences between drugs within and across the IL-17 and IL-23 inhibitor classes are concerned, several NMAs have shown the efficacy of ixekizumab and brodalumab to be

significantly greater than secukinumab, and similar to guselkumab and risankizumab.

Safety

Safety outcomes were analyzed in seven NMAs [17, 18, 21, 26, 33, 34, 38]. Five NMAs investigated the proportion experiencing any adverse event (AE), five investigated the incidence of serious AEs, and two others looked at infectious AEs and discontinuations due to AEs, respectively. Due to the small numbers of NMAs assessing safety outcomes and variation in the interventions considered, comparisons between these analyses were very limited. Full results can be seen in the supplementary materials.

DISCUSSION

Findings in Context

The availability of multiple biologic treatments for moderate-to-severe plaque psoriasis has resulted in their efficacy and safety being compared in numerous NMAs. Although NMAs are often used to support reimbursement decisions [56], doubts regarding their credibility may have limited their application to clinical decision making [12]. This SLR is, to our knowledge, the first to review all published NMAs evaluating biologics for the treatment of psoriasis. We have included 25 NMAs that were published between 2006 and 2020.

We found that 23 NMAs provided results for different levels of PASI response, and only seven provided safety results. Safety results were often inconclusive and inconsistent across analyses and varied in terms of the outcomes considered. Although an important aspect of treatment comparisons, safety appears to be disregarded or insufficiently analyzed.

We found that, overall, the included NMAs met at least half of the criteria included in the ISPOR checklist, reflecting the moderate to good methodological quality of the NMAs. The majority of NMAs appeared to take sufficient steps to identify all relevant trials, which always formed connected networks. All used methods

Table 4 Efficacy ranking of treatments at PASI 90 or PASI75 across network meta-analyses

NMA	Ranking; outcome	IFX	ETN 50	ETN 100	ETN any	ADA	UST 90	UST 45	UST WB/any	SEC	IXE	BRO	GUS	TIL	RIS	CZP
Woolcott 2006	NR; PASI 90	1 (NR)			2 (NR)											
Reich 2008	NR; PASI 90	1 (NR)	3 (NR)	2 (NR)												
Bansback 2009	NR; PASI 90	1 (NR)		3 (NR)		2 (NR)										
Reich 2012	Yes; PASI 90	1 (NR)			5 (NR)	4 (NR)	2 (NR)	3 (NR)								
Lin 2012	Yes; PASI 90	1 (ADA, ETN, UST)			4 (none)	3 (none)	2 (ADA, ETN)									
Gupta 2014	Yes; PASI 75	1 (UST, ADA, ETN)			2 (none)	2 (none)	2 (ETN) ^a									
Signorovitch 2015	NR; PASI 75 ^b	1 (ADA, UST, ETN)	6 (none)	5 (ETN 50)		4 (ETN)	2 (ETN)	3 (ETN)								
Gomez Garcia 2016	Yes; PASI 90	2 (ADA, ETN)	8 (none)	7 (ETN 50)		6 (none)	3 (ETN)	4 (ETN)	5 (ETN 50)	1 (ADA, ETN, UST)						
Jabbar-Lopez 2017	Yes; PASI90 ^c	3 (ETN)			6 (none)	5 (none)			4 (ETN)	2 (ADA, ETN, UST)	1 (ADA, ETN, IFX, UST)					
Sbidian 2017	Yes; PASI 90	9 (none)			10 (none)	8 (none)			6 (ETN)	2 (ADA, ETN, IFX, UST)	1 (ADA, ETN, IFX, UST)	3 (ADA, ETN, IFX, UST)	4 (ADA, ETN, IFX)	7 (none)		5 (none)
Geng 2018	Yes; PASI 90	2 (ETN 50)	6 (none)	5 (ETN 50)		4 (none)	1 (ETN)	3 (ETN)								
Loos 2018	NR; PASI 90	3 (UST, ADA, ETN)				6 (none)			5 (ETN)	4 (ETN)	1 (SEC, UST, ADA, ETN)	2 (UST, ADA, ETN)				
Sawyer 2018 (i)	NR; PASI 90	4 (NR)	8 (NR)				5 (NR)	6 (NR)	7 (NR)	3 (NR)	1 (NR)	2 (ADA, ETN, IFX, SEC, UST)				
Sawyer 2018 (m)	Yes; PASI 90	6 (ETN)		9 (none)		8 (none)	4 (ETN)	5 (ETN)	7 (ETN)	3 (UST, ETN)	2 (UST, ETN)	1 (ADA, ETN, IFX, SEC, UST)				
Cameron 2018	Yes; PASI 90	5 (NR)	12 (NR)	11 (NR)		6 (NR)	7 (NR)	8 (NR)	9 (NR)	4 (NR)	2 (NR)	3 (NR)	1 (SEC, IFX, ADA, UST, TIL, ETA)	10 (NR)		
Sawyer 2019	NR; PASI 90	6 (ADA, CZP, ETN, INF, TIL, UST)	12 (none)	11 (none)		7 (ETN)			8 (ETN)	5 (ADA, CZP, ETN, INF, TIL, UST)	1 (ADA, CZP, ETN, INF, SEC, TIL, UST)	3 (ADA, CZP, ETN, INF, SEC, TIL, UST)	4 (ADA, CZP, ETN, INF, TIL, UST)	10 (ETN)	2 (ADA, CZP, ETN, INF, SEC, TIL, UST)	9 (ETN)
Bai 2019	Yes; PASI 75						5 (UST 45)	8 (none)		2 (GUS, UST)	1 (GUS, UST)	3 (UST)	7 (none)	6 (none)	4 (UST)	
Xu 2019a	Yes; PASI 90	3 (ETN)			9 (none)	8 (none)			6 (ETN)	4 (ADA, ETN, TIL, UST)	2 (ADA, ETN, TIL, UST)	1 (ADA, ETN, GUS, TIL, UST)	5 (ADA, ETN, TIL, UST)	7 (ETN)		
Xu 2019b	Yes; PASI 75				3 (none)				5 (none)				1 (ADA, TIL, UST)	4 (none)	2 (UST)	
Warren 2019	NR; PASI 90 ^d	3 (NR)		10 (NR)		6 (NR)	7 (NR)	8 (NR)	7 (NR)	4 (NR)	2 (NR)	1 (NR)	5 (NR)	9 (NR)		
Sbidian 2020	Yes; PASI 90	1 (ADA, CZP, ETN, TIL, UST)			11 (none)	7 (CZP, ETN)			8 (CZP, ETN)	5 (ADA, CZP, ETN, TIL, UST)	2 (ADA, BRO, CZP, ETN, SEC, TIL, UST)	6 (ADA, CZP, ETN, UST)	4 (ADA, CZP, ETN, TIL, UST)	9 (ETN)	3 (ADA, BRO, CZP, ETN, TIL, UST)	10 (none)
Warren 2020	Yes; PASI 90 ^e	5 (NR)		11 (NR)		8 (NR)			7 (NR)	4 (NR)	1 (NR)	3 (NR)	6 (NR)	10 (NR)	2 (NR)	9 (NR)
Armstrong 2020	NR; PASI 90	6 (NR)	11 (NR)			9 (NR)			8 (NR)	5 (NR)	2 (NR)	3 (NR)	4 (NR)	10 (NR)	1 (NR)	7 (NR)

Table 4 continued

ADA adalimumab, *BRO* brodalumab, *CZP* certolizumab pegol, *ETN* etanercept, *GUS* guselkumab, *IFX* infliximab, *IXE* ixekizumab, *N* not reported, *PASI* psoriasis area severity index, *RIS* risankizumab, *SEC* secukinumab, *TIL* tildrakizumab, *UST* ustekinumab, *WB* weight-based dosing, *Y* reported. Green denotes treatments ranked 1–3, orange 4–6, and grey 7–9. Treatments in brackets are statistically significantly less efficacious

^a Paper considered this approximately the same rank (reported “ustekinumab \approx adalimumab \approx etanercept” based on PASI 75 in spite of results OR for ustekinumab vs. etanercept 1.94 (95% CrI 1.31, 3.01)

^b The multinomial probit model used in this analysis ensures that the rankings observed at PASI 75 will apply for PASI 90

^c Reported as “clear/nearly clear” but input data related to PASI 90 or PGA 0/1

^d Based on maximum area under the curve

^e Based on outcomes at week 12

that preserve within-study randomization and reported the evidence network in graphical or tabular form. One of the major limitations of the identified analyses was that only half of the NMAs discussed the rationale for the choice of the model. In addition, just over half of analyses addressed the impact of bias due to differences in effect modifiers, through sensitivity analyses and use of different models, for example [17, 24, 26, 30, 31, 33, 34, 36–38, 40, 42, 43].

Twenty-one NMAs provided details of their funding and conflicts of interest. After accounting for uncertainty in the head-to-head treatment effects and relative ranks, all analyses came to similar conclusions, regardless of funding source. Publicly funded and industry funded analyses that used similar methods reported similar results and conclusions. Any apparent differences by funding source are better explained by comparing the methodological approach and assumptions underpinning the synthesis of evidence.

The focus of our analysis was the comparisons of the treatment efficacy compared with placebo for PASI outcomes at the end of the induction phase, as these were most widely reported. This comparison did not include Sawyer 2018 (m) [40], as it was the only analysis to investigate longer-term efficacy and was therefore not comparable with other NMAs. In addition, Gomez-Garcia 2016 [23] did not report results for active treatments compared with placebo and therefore could not be considered here either.

We identified that the efficacy of older biologics (anti-TNF agents) compared with placebo appeared to decrease slightly from

approximately 2015 onwards. One potential explanation for this trend could be a change in the percentage of patients in trials who are biologic experienced, due to biologic therapies becoming the standard of care. If more recent trials included a higher proportion of patients who have previously tried TNF inhibitors, the proportion of patients responding to these treatments could be lower, compared with earlier trials. It also has the potential to affect the placebo response, which would therefore support the idea that analyses must be placebo-adjusted.

When analyses published within a similar time period were considered, their results were broadly similar, emphasizing the consistency in reported treatment effects. The only major exception appeared to be the original and updated Cochrane Reviews [33, 34], which consistently provided substantially lower estimates of treatment efficacy compared with placebo than other NMAs published around the same time. These NMAs were outliers in some methodological respects, including the choice of approach (frequentist, rather than Bayesian), the handling of PASI outcomes (dichotomous), and evaluated doses.

Six analyses used a frequentist approach and eighteen analyses used a Bayesian one, with one study employing both methods and demonstrating the statistical approach to have no impact on the results [20]. Treatment effects are reported in a number of different ways, including risk ratios, odds ratios, risk differences, and numbers needed to treat, and this variety makes it difficult to compare results

across studies beyond directional trends and conclusions of statistical difference.

Considering model choice, we found that NMAs either analyzed every level of PASI response separately (binomial models) or jointly (multinomial models). Whilst the choice of model only appears to result in small differences between treatment efficacy and safety estimates, these can potentially lead to different analysis results when considering the relative ranking of treatments. The use of a multinomial model allows inclusion of all available information. It also has the advantage of using the evidence on the relationships between different PASI response levels to make predictions about the relative performance of a treatment at a threshold not reported in a given study or studies, thus allowing comparisons that would be otherwise impossible using a binomial model [5]. This approach results in coherence across all PASI outcomes, a feature crucial when the results of an analysis are utilized in an economic model. However, when the major purpose of the analysis is to provide the best estimate of treatment efficacy at a specific level of response, a binomial approach may be more helpful. One of the potential advantages of using a binomial approach is that it only includes evidence from studies that reported results for a particular level of PASI response and therefore avoids the additional uncertainty resulting from extrapolation from other PASI levels. It also allows for more subtle differences between the responses at different levels of PASI to be observed. However, it appears that in the case of psoriasis the choice of binomial or multinomial model did not substantially affect estimates of efficacy. This was clearly highlighted in Cameron 2018 [38], where the base-case analysis was performed using a binomial model and a supplementary analysis was carried out using a multinomial model. Based on the results reported, relatively small differences between both analyses in treatment efficacy were seen.

For their base-case analysis, 17 NMAs used a random-effects model, three used a fixed-effect model and five did not state what model was used. Placebo adjustment was reported in six NMAs [24, 30, 36–38, 42]. This method accounts for cross-trial differences that may

otherwise be difficult to identify and incorporate in the model. It will also likely show if the analysis results are influenced by these differences, thus minimizing inaccuracies in analysis results [57]. All six NMAs found the adjusted model was a better fit for the data than the unadjusted one.

As can be seen in the online supplement, the trials included in NMAs were broadly consistent, taking into account the time of publication. However, there were some cases where the inclusion criteria of the analyses differed, potentially leading to slight differences in results.

One of the reasons for the discrepancies in the trials included in the NMAs was the time-point of interest for the analyses. Across analyses, the timepoints at which treatments were compared varied between 8 and 24 weeks. One of the major reasons for discrepancies appears to be the inclusion or exclusion of the 10-week timepoint corresponding to the duration of the majority of infliximab trials. Only the 2017 Cochrane Review [34] did not include 10-week data. As a result, this NMA excluded some major infliximab trials [58–60] and based its efficacy estimate on two trials comparing infliximab with methotrexate and etanercept [61, 62]. It reported lower estimates of infliximab efficacy compared with other analyses. Although including only trials with a small range of follow-up times may reduce heterogeneity, it may not always result in a clinically relevant estimate. As biologics used in psoriasis have a different induction period at which point a decision is made on whether to continue treatment, it appears that providing estimates relevant to that decision point (which varies from 10 to 16 weeks) may be of more importance. It also ensures the inclusion of the most relevant evidence base.

In some cases, different doses of treatments for psoriasis are included in clinical trials. In particular, phase 2 studies may include a range of doses later judged as subtherapeutic or intolerable and not pursued in later-phase research. The identified NMAs that included a range of treatment doses have handled these differently.

Whilst some analyzed different treatment doses separately, others pooled evidence on some or all doses of a drug [17, 25, 33, 34]. Although pooling doses may provide the advantage of utilizing all available evidence, analyzing doses separately is generally more relevant to clinical decision making, because in clinical practice patients are prescribed the licensed dose of a drug. The efficacy and safety of drugs is usually dose-dependent and therefore the effect estimates obtained by pooling data for unlicensed and licensed doses may not provide a reliable estimate of the treatment effects. In the case of biologics in psoriasis it is likely to underestimate the efficacy of a treatment, as the doses which have not been marketed are frequently lower and do not offer the full therapeutic benefit. Although the evidence on safety was limited, pooling of licensed and unlicensed doses may also result in the under or overestimation of adverse events.

The impact of pooling evidence for different doses was most pronounced in the base-case analysis of the Cochrane Reviews [33, 34]. In contrast, the results of Jabbar-Lopez 2017 [17] appeared only slightly less favorable than in the other NMAs of that time period, likely due to the inclusion of few treatments with evidence for multiple dosing regimens. Nevertheless, for clinical decision making, these results may not provide the best summary of the relevant evidence base.

However, it must be noted that whilst both the Cochrane Reviews and Jabbar-Lopez 2017 pooled doses for their base-case analysis, they did also perform sensitivity analyses (available in supplementary analyses) in which doses were not pooled. It is, therefore, more appropriate to say that the choice of base-case should be more carefully considered.

Limitations

There were several limitations to this review. Firstly, we did not include HTA reports. Although these are of high importance and largely shape clinical practice in different countries, the details of these analyses are often confidential and therefore their inclusion in our

analysis would likely not be possible. However, some of these have been published and were therefore included in our analysis [24, 35]. For similar reasons we did not consider conference abstracts to be eligible in this SLR. When screening studies, the second reviewer only performed a 40% check, resulting in the potential for missing some NMAs. This was unlikely, however, as the level of disagreement between the two reviewers was low and the chances of missing studies were further minimized by checking reference lists in identified publications. Also, we only included analyses in English; however, this is likely to have excluded few, if any, relevant analyses.

Only NMAs that included two or more biologics were of interest for our analysis. Whilst non-biologic systemic treatments are used in moderate-to-severe psoriasis and are often licensed in the same population, in practice they would be used before biologics. Therefore, we believe that their inclusion in our analysis would add unnecessary complexity. Similarly, we excluded simple and matched indirect treatment comparisons, where only two treatments were compared with placebo. In an area such as psoriasis, with its abundance of therapeutic options, such analyses would only provide a very narrow view of the treatment landscape.

Furthermore, we compared results from NMAs for PASI response, as these were the most commonly reported results. It is possible, however, that if we considered outcomes such as Physician's Global Assessment (PGA) and Dermatology Life Quality Index (DLQI), the results of our analysis could have been different. The scope for comparison of the included NMA evidence for these outcomes was limited, however, as only five analyses evaluated PGA, [21, 25, 33, 34, 38] and five, DLQI [17, 20, 25, 33, 34].

CONCLUSIONS

There are numerous analyses assessing the efficacy and safety of treatments for moderate-to-severe psoriasis. Despite their methodological differences and differences in funding sources

the analyses reported broadly similar results and conclusions. This consistency highlights the reliability of NMAs for use in clinical practice, emphasizing that newer biologics are more efficacious than older treatments. However, there were some important differences in the results of NMAs, likely resulting from the methods used and assumptions made.

When using NMAs to inform clinical practice, consideration should be given to those NMAs that include most, if not all, of the features that make a methodologically valid and robust analysis: inclusion of all relevant trials and comparator treatments, thorough assessment of heterogeneity and inconsistency, complete reporting of comparative effects, and associated uncertainty. To ensure that results are relevant and applicable, comparator treatments should be synthesized in a way that is reflective of their marketing authorization and use in clinical practice, considering characteristics such as population, dose, and duration. Finally, whilst efficacy is of high importance to patients, so are outcomes such as quality of life and safety. An analysis that captures all of these factors would be optimal for use in clinical decision making.

ACKNOWLEDGEMENTS

Funding. This study and the journal's Rapid Service Fee were funded by LEO Pharma A/S.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published. The authors would like to thank Jana Tillotson for graphic design support during the preparation of this article.

Disclosures. Najeeda Yasmeen and Laura Sawyer are employed by Symmetron Limited, which received funding from LEO Pharma for this research. At the time the research was undertaken, Emily Wright and Kinga Malottki

were also employed by Symmetron Limited. Kinga Malottki is currently an employee of Sanofi. Emma Borg is an employee of LEO Pharma. Carsten Schwenke is self-employed and received funding from LEO Pharma for this research. Richard Warren has received research grants and/or consulting fees from AbbVie, Almirall, Amgen, Arena, Avillion, Bristol-Myers Squibb, Boehringer Ingelheim, Celgene, Eli Lilly, Janssen, LEO Pharma, Novartis, Pfizer, Sanofi, and UCB Pharma. Richard Warren is supported by Manchester National Institute for Health Research Biomedical Research Centre. Richard Warren is also a member of the journal's Editorial Board.

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

Data Availability. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

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. Goff K, Karimkhani C, Boyers L, et al. The global burden of psoriatic skin disease. *Br J Dermatol*. 2015;172(6):1665–8.
2. Weger W. Current status and new developments in the treatment of psoriasis and psoriatic arthritis with biological agents. *Br J Pharmacol*. 2010;160(4):810–20.
3. Ogawa E, Sato Y, Minagawa A, Okuyama R. Pathogenesis of psoriasis and development of treatment. *J Dermatol*. 2018;45(3):264–72.
4. Tonini A, Gualtieri B, Panduri S, Romanelli M, Chiricozzi A. A new class of biologic agents facing the therapeutic paradigm in psoriasis: anti-IL-23 agents. *Expert Opin Biol Ther*. 2018;18(2):135–48.
5. 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*. 2013;33(5):607–17.
6. Li T, Puhan MA, Vedula SS, Singh S, Dickersin K, Group TAHNM-aMMW. Network meta-analysis—highly attractive but more methodological research is needed. *BMC Med*. 2011;9(1):79.
7. Tonin FS, Rotta I, Mendes AM, Pontarolo R. Network meta-analysis: a technique to gather evidence from direct and indirect comparisons. *Pharm Pract (Granada)*. 2017;15(1):943.
8. Kibret T, Richer D, Beyene J. Bias in identification of the best treatment in a Bayesian network meta-analysis for binary outcome: a simulation study. *Clin Epidemiol*. 2014;6:451–60.
9. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. London: National Institute for Health and Care Excellence; 2013.
10. Jansen JP, Fleurence R, Devine B, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health*. 2011;14(4):417–28.
11. Bhatnagar N, Lakshmi P, Jeyashree K. Multiple treatment and indirect treatment comparisons: an overview of network meta-analysis. *Perspect Clin Res*. 2014;5(4):154.
12. Lee AW. Use of network meta-analysis in systematic reviews: a survey of authors. *Syst Rev*. 2016;5(1):8.
13. 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(6):683–91.
14. Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health*. 2014;17(2):157–73.
15. Ades A, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S. NICE DSU Technical Support Document 7: evidence synthesis of treatment efficacy in decision making: a reviewer's checklist. The Decision Support Unit London UK: National Institute for Health and Care Excellence. 2012.
16. R Core Team. R: A language and environment for statistical computing. Vienna, Austria. 2020.
17. Jabbar-Lopez ZK, Yiu ZZN, Ward V, Exton LS, Mohd Mustapa MF, Samarasekera E, Burden AD, Murphy R, Owen CM, Parslew R, Venning V, Warren RB, Smith CH. Quantitative evaluation of biologic therapy options for psoriasis: a systematic review and network meta-analysis. *J Investig Dermatol*. 2017;137(8):1646–54.
18. Messori A, Trippoli S, Fadda V, Maratea D, Marinai C. Subcutaneous biological treatments for moderate to severe psoriasis: interpreting safety data by network meta-analysis. *Drugs Real World Outcomes*. 2015;2(1):23–7.
19. Gupta AK, Daigle D, Lyons DCA. Network meta-analysis of treatments for chronic plaque psoriasis in Canada. *J Cutaneous Med Surg*. 2014;18(6):371–8.
20. 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*. 2020;10(1):73–86.
21. Bai F, Li GG, Liu Q, Niu X, Li R, Ma H. 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.
22. Geng W, Zhao J, Fu J, Zhang H, Qiao S. Efficacy of several biological therapies for treating moderate to severe psoriasis: a network meta-analysis. *Exp Ther Med*. 2018;16(6):5085–95.

23. Gómez-García F, Epstein D, Isla-Tejera B, Lorente A, Vélez García-Nieto A, Ruano J. 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(3):594–603.
24. Loos AM, Liu S, Segel C, Ollendorf DA, Pearson SD, Linder JA. 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(1):135e7-144.
25. 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(4):289–303.
26. Xu S, Zhang X, Pan M, Shuai Z, Xu S, Pan F. Treatment of plaque psoriasis with IL-23p19 blockers: a systematic review and meta-analysis. *Int Immunopharmacol*. 2019;75:105841.
27. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996;17(1):1–12.
28. Higgins JP, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
29. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Updated 7 June 2020. https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed 13 Mar 2020.
30. 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.
31. Bansback N, Sizto S, Sun H, Feldman S, Willian MK, Anis A. Efficacy of systemic treatments for moderate to severe plaque psoriasis: systematic review and meta-analysis. *Dermatology*. 2009;219(3):209–18.
32. Reich K, Sinclair R, Roberts G, Griffiths CEM, Tabberer M, Barker J. Comparative effects of biological therapies on the severity of skin symptoms and health-related quality of life in patients with plaque-type psoriasis: a meta-analysis. *Curr Med Res Opin*. 2008;24(5):1237–54.
33. 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.
34. Sbidian E, Chaimani A, Garcia-Doval I, et al. Systemic pharmacological treatments for chronic plaque psoriasis: a network meta-analysis. *Cochrane Library*. 2017.
35. Woolacott N, Hawkins N, Mason A, et al. Efalizumab and Etanercept for the Treatment of Psoriasis. CRD Technology Assessment Group University of York on behalf of NICE. 2005:1–356.
36. Sawyer L, Fotheringham I, Wright E, Yasmeeen N, Gibbons C, Holmen Møller A. 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.
37. 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 [Electronic Resource]*. 2019;14(8):e0220868.
38. Cameron C, Druchok C, Hutton B, et al. Guselkumab for the treatment of moderate-to-severe plaque psoriasis during induction phase: a systematic review and network meta-analysis. *J Psoriasis Psoriatic Arthritis*. 2018;4:81–92.
39. 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(1):179–88.
40. Sawyer LM, Cornic L, Levin LA, Gibbons C, Moller AH, Jemec GB. 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(2):355–66.
41. 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(5):1138–49.
42. 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(2):504–12.
43. 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(12):1403–10.

44. Wu MY, Yu CL, Yang SJ, Chi CC. Change in body weight and body mass index in psoriasis patients receiving biologics: a systematic review and network meta-analysis. *J Am Acad Dermatol.* 2020;82(1):101–9.
45. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol.* 2011;64(2):163–71.
46. Elewski BE, Okun MM, Papp K, et al. Adalimumab for nail psoriasis: efficacy and safety from the first 26 weeks of a phase 3, randomized, placebo-controlled trial. *J Am Acad Dermatol.* 2018;78(1):90e1–99.
47. Reich K, Sullivan J, Arenberger P, et al. Effect of secukinumab on the clinical activity and disease burden of nail psoriasis: 32-week results from the randomized placebo-controlled TRANSFIGURE trial. *Br J Dermatol.* 2019;181(5):954–66.
48. Mease PJ, van der Heijde D, Ritchlin CT, et al. Ixekizumab, an interleukin-17A specific monoclonal antibody, for the treatment of biologic-naïve patients with active psoriatic arthritis: results from the 24-week randomised, double-blind, placebo-controlled and active (adalimumab)-controlled period of the phase III trial SPIRIT-P1. *Ann Rheum Dis.* 2017;76(1):79–87.
49. Sterry W, Ortonne JP, Kirkham B, et al. Comparison of two etanercept regimens for treatment of psoriasis and psoriatic arthritis: PRESTA randomised double blind multicentre trial. *BMJ.* 2010;340:c147.
50. Leonardi C, Matheson R, Zachariae C, et al. Anti-interleukin-17 monoclonal antibody ixekizumab in chronic plaque psoriasis. *N Engl J Med.* 2012;366(13):1190–9.
51. Papp KA, Langley RG, Sigurgeirsson B, et al. Efficacy and safety of secukinumab in the treatment of moderate-to-severe plaque psoriasis: a randomized, double-blind, placebo-controlled phase II dose-ranging study. *Br J Dermatol.* 2013;168(2):412–21.
52. Rich P, Sigurgeirsson B, Thaci D, et al. Secukinumab induction and maintenance therapy in moderate-to-severe plaque psoriasis: a randomized, double-blind, placebo-controlled, phase II regimen-finding study. *Br J Dermatol.* 2013;168(2):402–11.
53. Papp KA, Blauvelt A, Bukhalo M, et al. Risankizumab versus ustekinumab for moderate-to-severe plaque psoriasis. *N Engl J Med.* 2017;376(16):1551–60.
54. Nemoto O, Hirose K, Shibata S, Li K, Kubo H. Safety and efficacy of guselkumab in Japanese patients with moderate-to-severe plaque psoriasis: a randomized, placebo-controlled, ascending-dose study. *Br J Dermatol.* 2018;178(3):689–96.
55. Sofen H, Smith S, Matheson RT, et al. Guselkumab (an IL-23-specific mAb) demonstrates clinical and molecular response in patients with moderate-to-severe psoriasis. *J Allergy Clin Immunol.* 2014;133(4):1032–40.
56. Maynou L, Cairns J. What is driving HTA decision-making? Evidence from cancer drug reimbursement decisions from 6 European countries. *Health Policy.* 2019;123(2):130–9.
57. 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(11):1037–51.
58. Gottlieb AB, Evans R, Li S, et al. Infliximab induction therapy for patients with severe plaque-type psoriasis: a randomized, double-blind, placebo-controlled trial. *J Am Acad Dermatol.* 2004;51(4):534–42.
59. Reich K, Nestle FO, Papp K, et al. Infliximab induction and maintenance therapy for moderate-to-severe psoriasis: a phase III, multicentre, double-blind trial. *Lancet.* 2005;366(9494):1367–74.
60. Menter A, Feldman SR, Weinstein GD, et al. A randomized comparison of continuous vs. intermittent infliximab maintenance regimens over 1 year in the treatment of moderate-to-severe plaque psoriasis. *J Am Acad Dermatol.* 2007;56(1):31.
61. Barker J, Hoffmann M, Wozel G, et al. Efficacy and safety of infliximab vs. methotrexate in patients with moderate-to-severe plaque psoriasis: results of an open-label, active-controlled, randomized trial (RESTORE1). *Br J Dermatol.* 2011;165(5):1109–17.
62. De Vries A, Thio H, de Kort W, et al. A prospective randomized controlled trial comparing infliximab and etanercept in patients with moderate-to-severe chronic plaque-type psoriasis: the Psoriasis Infliximab vs. Etanercept Comparison Evaluation (PIECE) study. *Br J Dermatol.* 2017;176(3):624–33.