



Fahrbach, K., Tarpey, J., Bergrath, E., Hughes, R., Thom, H. H. Z., Neary, M., Cha, A., Gerber, R., & Cappelleri, J. (2020). Crisaborole Ointment, 2%, for Treatment of Patients with Mild-to-Moderate Atopic Dermatitis: Systematic Literature Review and Network Meta-Analysis. *Dermatology and Therapy*, *10*, 681–694.  
<https://doi.org/10.1007/s13555-020-00389-5>

Publisher's PDF, also known as Version of record

License (if available):  
CC BY-NC

Link to published version (if available):  
[10.1007/s13555-020-00389-5](https://doi.org/10.1007/s13555-020-00389-5)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the final published version of the article (version of record). It first appeared online via Springer Nature at [10.1007/s13555-020-00389-5](https://doi.org/10.1007/s13555-020-00389-5). Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



# Crisaborole Ointment, 2%, for Treatment of Patients with Mild-to-Moderate Atopic Dermatitis: Systematic Literature Review and Network Meta-Analysis

Kyle Fahrbach · Jialu Tarpey · Evelien Bergrath Washington ·  
Rachel Hughes · Howard Thom · Maureen P. Neary · Amy Cha ·  
Robert Gerber · Joseph C. Cappelleri

Received: April 3, 2020 / Published online: May 20, 2020  
© The Author(s) 2020

## ABSTRACT

**Introduction:** There is a need to compare efficacy and safety profiles of crisaborole ointment, 2%, versus other topical treatments across randomized clinical trials (RCTs). We performed this review/network meta-analysis to evaluate the comparative efficacy and safety of crisaborole versus other topical pharmacologic therapies for mild-to-moderate atopic dermatitis (AD) among patients aged  $\geq 2$  years.

**Digital Features** To view digital features for this article go to <https://doi.org/10.6084/m9.figshare.12199766>.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s13555-020-00389-5>) contains supplementary material, which is available to authorized users.

K. Fahrbach · J. Tarpey · E. B. Washington (✉) ·  
R. Hughes  
Evidence Synthesis, Modeling & Communication,  
Evidera, Waltham, MA, USA  
e-mail: [Evelien.bergrathwashington@evidera.com](mailto:Evelien.bergrathwashington@evidera.com)

H. Thom  
Population Health Sciences, Bristol Medical School,  
University of Bristol, Bristol, UK

H. Thom  
Clifton Insight, Bristol, UK

M. P. Neary  
Inflammation and Immunology, Pfizer Inc.,  
Collegeville, PA, USA

**Methods:** Searches were conducted in MEDLINE, Embase, the Cochrane Collection Central Register of Clinical Trials, and the Database of Abstracts of Reviews of Effects using Ovid to identify English language articles reporting RCTs of topical anti-inflammatory agents in patients aged  $\geq 2$  years with mild-to-moderate AD published between inception and 10 March 2020. This review used a prespecified protocol with eligibility criteria for population, interventions, comparisons, outcomes, and study design. Efficacy was evaluated using the Investigator's Static Global Assessment (ISGA) of clear (0) or almost clear (1) and expressed by hazard ratios (HR) with 95% credible intervals.

**Results:** Patients treated with crisaborole or tacrolimus ointment, 0.1% or 0.03%, versus vehicle alone were significantly more likely to achieve ISGA 0/1 at 28–42 days, with the

A. Cha  
Inflammation and Immunology, Pfizer Inc., New  
York, NY, USA

R. Gerber  
Inflammation and Immunology, Pfizer Inc., Groton,  
CT, USA

J. C. Cappelleri  
Global Biometrics and Data Management  
(Statistics), Pfizer Inc., Groton, USA

greatest point estimate observed for the crisaborole comparison (hazard ratio: 2.07; 95% credible interval 1.76 to – 2.36; probability HR above 1 [*p* better]: 100.0%). Patients were also more likely to achieve ISGA 0/1 with crisaborole than with pimecrolimus cream, 1% (HR: 1.62; 95% credible interval 1.04–2.48; *p* better: 98.3%). While network meta-analysis for safety was not feasible because of data limitations, crisaborole pivotal studies (AD-301/AD-302) showed crisaborole was well tolerated.

**Conclusions:** Crisaborole was shown to be superior to vehicle and pimecrolimus and comparable to tacrolimus, 0.1% or 0.03%, with respect to ISGA 0/1 at 28–42 days in patients aged  $\geq 2$  years with mild-to-moderate AD. This evaluation of comparative efficacy of crisaborole further supports use of crisaborole as an effective therapeutic option in this population.

**Keywords:** Atopic dermatitis; Crisaborole; Network meta-analysis; Systematic literature review

### Key Summary Points

The objective of this systematic literature review (SLR) and network meta-analysis (NMA) was to evaluate the comparative efficacy and safety of crisaborole versus other topical pharmacologic treatments for mild-to-moderate atopic dermatitis (AD) among patients aged  $\geq 2$  years

Our search of Embase<sup>®</sup>, MEDLINE<sup>®</sup>, CENTRAL, and DARE using Ovid identified 894 articles published through 10 March 2020; after screening and the feasibility assessment, nine unique randomized clinical trials (RCTs) were deemed eligible for evaluation through NMA

Efficacy was evaluated using the Investigator's Static Global Assessment (ISGA) score of clear (0) or almost clear (1) at 28–42 days with relative treatment effects expressed by hazard ratios (HR) with 95% credible intervals

Crisaborole 2% ointment was shown to be superior to pimecrolimus 1% cream and vehicle, and comparable to tacrolimus 0.1% or 0.03% ointment, in achieving the ISGA 0/1 score at 28–42 days

This evaluation of comparative efficacy of crisaborole 2% ointment further supports its use as an effective therapeutic option in patients aged  $\geq 2$  years with mild-to-moderate AD

## INTRODUCTION

Atopic dermatitis (AD), a common chronic inflammatory skin disorder characterized by eczematous, lichenified lesions and intense pruritus, usually appears in childhood and is often associated with comorbidities such as asthma and allergic rhinitis [1–3]. AD affects 15–20% of children (< 18 years of age) and 1–3% of adults [4]. Most patients with AD suffer from mild-to-moderate disease [5–7].

The goal of AD management is the prevention and care of disease flares. US treatment guidelines recommend topical corticosteroids (TCSs) and/or topical calcineurin inhibitors (TCIs) as well as phototherapy for mild-to-moderate AD and immunosuppressants or biologics for moderate-to-severe/refractory disease [8]. Although there are safety concerns with the prolonged use of high-potency TCSs [9], a more significant problem is nonadherence to therapy because of fear of skin atrophy, which in turn leads to poor disease control [10]. While TCIs reduce AD severity, special warnings highlight a possible risk for lymphoma and skin cancer, and application site reactions may reduce its use [3, 8, 11].

Crisaborole is a nonsteroidal topical phosphodiesterase 4 inhibitor (PDE4i) that acts by regulating inflammatory cytokine production, which is overactive in patients with AD [12, 13]. Crisaborole was initially approved by the US Food and Drug Administration (FDA) in December 2016 for use as a topical treatment of mild-to-moderate atopic dermatitis in patients  $\geq 2$  years of age. In March 2020, the FDA

approved a supplemental New Drug Application that expanded the use of crisaborole to include children  $\geq 3$  months of age. Crisaborole was approved in the European Union in March 2020 for the treatment of mild-to-moderate atopic dermatitis in adults and pediatric patients from 2 years of age with  $\leq 40\%$  body surface area affected.

Crisaborole was previously approved in Australia, Canada, and Israel. Crisaborole applied twice daily was shown to be effective in patients  $\geq 2$  years of age with mild-to-moderate AD and was associated with a low incidence of treatment-related/treatment-emergent adverse events (AEs) [14]. A recent systematic review and network meta-analysis for PDE4is versus vehicle has shown that topical PDE4is are more effective than vehicle alone for patients with mild-to-moderate AD [15]. Nevertheless, there is a need to compare crisaborole with other topical treatments and to synthesize available evidence from newly published randomized clinical trials (RCTs).

A systematic literature review and a network meta-analysis were performed to evaluate the comparative efficacy and safety of crisaborole versus other topical pharmacologic therapies for mild-to-moderate AD among patients aged  $\geq 2$  years.

## METHODS

### Systematic Literature Review

Searches were conducted in MEDLINE (Ovid), Embase (Ovid), the Cochrane Collection Central Register of Clinical Trials (CENTRAL; Ovid), and the Database of Abstracts of Reviews of Effects (DARE; Ovid) to identify English language articles published between inception and 10 March 2020 reporting RCTs for evaluation of possible treatments for patients with mild-to-moderate AD. This systematic literature review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [16, 17]. The search strategies included a combination of controlled vocabulary terms as well as free-text search terms for disease and study designs of interest

(Supplement Tables S1–S3). In addition, we hand-searched abstracts from the 2015–2018 scientific meetings of the American Academy of Dermatology and the European Academy of Dermatology and Venereology, as well as bibliographies of included publications and systematic reviews identified in the search.

### Identification and Selection of Studies

The review was conducted using a prespecified protocol. Predefined eligibility criteria involved the use of the Population, Interventions, Comparisons, Outcomes, and Study design tool (PICOS; Table S4). Two blinded, independent reviewers examined the citations; any discrepancies were resolved by a third reviewer. The outcome of interest was Investigator's Static Global Assessment (ISGA) of 0/1 (clear/almost clear) at 28–42 days. Secondary outcomes of interest were AEs.

### Data Extraction

The relevant information extracted from eligible studies included study design and methods, patient characteristics, intervention details (e.g., dosing, schedule, components of vehicle), and efficacy and safety outcomes, along with time points for outcome assessments. A single reviewer extracted data, and a second reviewer quality-assessed the data accuracy.

### Quality Assessments

A risk-of-bias assessment was undertaken using the Cochrane tool, in accordance with the National Institute for Health and Care Excellence (NICE) single technology appraisal guidelines for evidence submissions [18, 19].

### Feasibility Assessments

Prior to analysis, a feasibility assessment determined the availability of evidence and identified potential sources of heterogeneity. All studies were compared with respect to study-

and patient-level characteristics, outcome definitions, and time points of evaluation.

### Network Meta-Analysis

A network meta-analysis was performed to obtain relative treatment effects for achievement of ISGA 0/1 at 28–42 days. All analyses were conducted within a Bayesian framework [20] and involved a 100,000-run-in iteration phase and a 100,000 iteration phase for parameter estimation. All calculations were performed using OpenBugs 3.2.3 [21]. Models using fixed effects and random effects on treatment effects were explored. Baseline risk regression was used to adjust for differences in vehicle response across RCTs; this was driven by variation in vehicle composition and by heterogeneity in patient characteristics. Baseline risk adjustment indirectly adjusted for heterogeneity in effect modifiers across RCTs [22]. Class-effects models with baseline risk regression used fixed effects across RCTs but random effects for treatments within class; classes included crisaborole, vehicle, and non-crisaborole treatments. Model fit was explored by comparing the deviance information criterion (DIC) and the posterior mean of the residual deviance for fixed- and random-effects models [23]. The model with the lowest DIC was considered to be the best fitting. Hazard ratios reflect the “hazard” of response; thus, hazard ratios (HRs) > 1.0 for comparisons between two treatments imply better performance for the first treatment. A detailed description of the statistical methods can be found in the Supplement.

### 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. The review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.

## RESULTS

### Systematic Literature Review

#### *Study Selection*

The search strategy identified 894 records after duplicates were removed, of which 212 were screened for full-text eligibility after title/abstract screening and duplicate removal. In total, nine RCTs (reported in 8 publications; 1 of which reported in 2 RCTs) were eligible for inclusion in the network meta-analysis (Fig. 1).

#### *Study Characteristics*

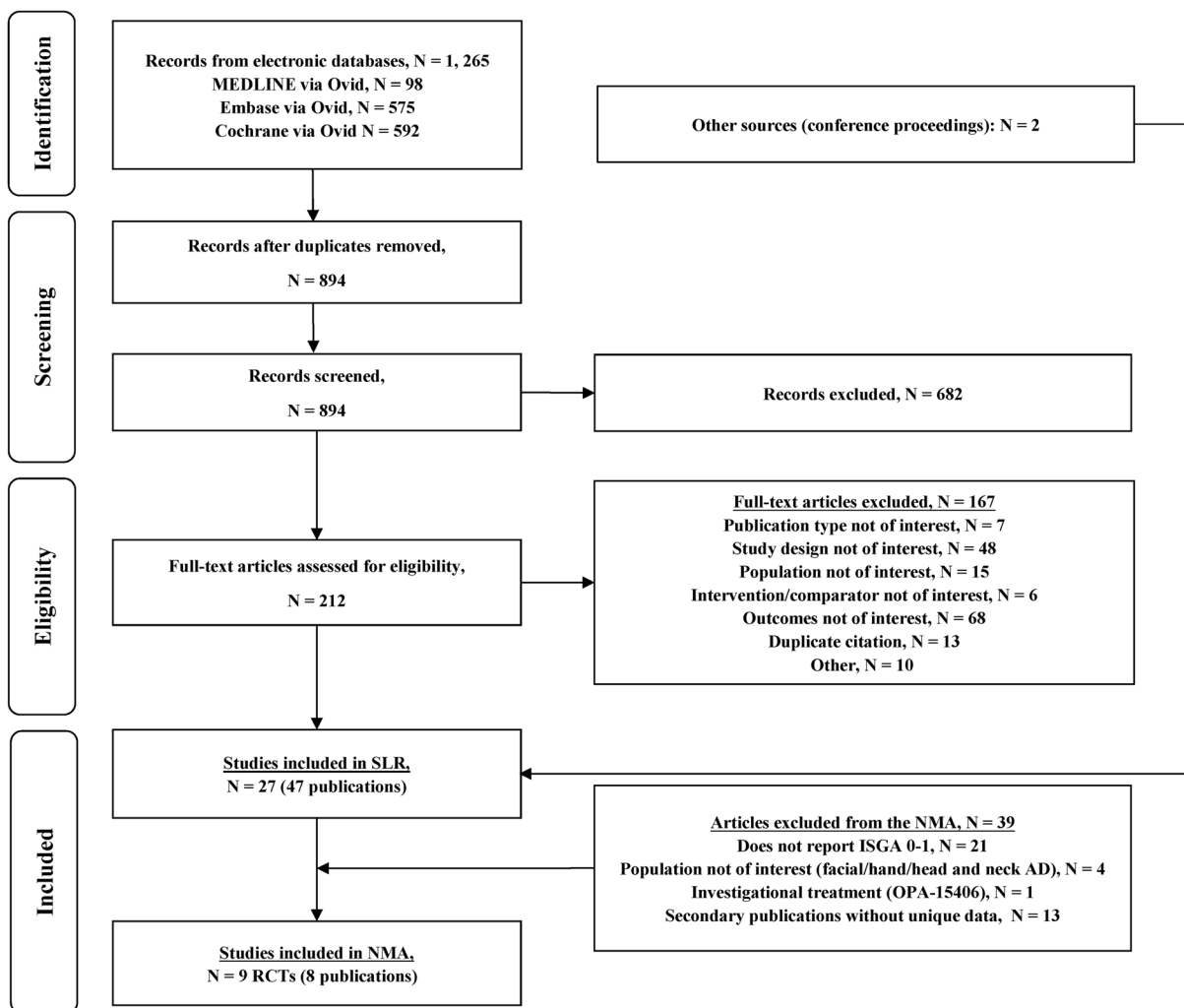
The follow-up duration across all included studies ranged from 28 [14] to 42 days [24]. Only five RCTs reported the recruitment period, which ranged from 2001–2015. Sample sizes varied from 133 [25] to 764 [14] (Table S5).

#### *Patient Characteristics*

The average age ranged from 6.4 [26] to 39.1 [27] years, reflecting a mix of pediatric and adult populations (one RCT did not report mean age [25]). Per the inclusion criteria, all evaluated patients  $\geq 2$  years of age, except for Eichenfield et al. [24], which enrolled patients aged 1–17 years and reported an overall average patient age of 6.7 years [24]. Most RCTs (44%) included children and adolescents (2–17 years); an additional 33% included adult and pediatric patients. Six RCTs enrolled mixed mild-to-moderate populations, whereas one enrolled only mild [26] and two enrolled only moderate [27, 28]. Most studies ( $n = 8$ ) defined disease severity according to ISGA. Three studies provided baseline Eczema Area and Severity Index scores, but none provided baseline SCORing of AD scores [24, 26, 27, 29]. Baseline percentage of body surface area affected measurements were provided by seven RCTs, ranging from 11.1% [30] to 25.9% [24].

#### *Treatment Characteristics*

Treatments evaluated included crisaborole (PDE4i), pimecrolimus, and tacrolimus, 0.1% or 0.03% (TCIs), each administered as monotherapy. All doses administered in the RCTs were FDA-approved for the treatment of AD.



**Fig. 1** PRISMA flow diagram of the systematic literature review. *AD* atopic dermatitis, *ISGA* Investigator’s Static Global Assessment, *NMA* network meta-analysis, *PRISMA*

Preferred Reporting Items for Systematic Reviews and Meta-Analyses; *RCT* randomized clinical trial, *SLR* systematic literature review

Pimecrolimus is approved for mild-to-moderate disease, whereas tacrolimus (in both available concentrations) is approved for moderate-to-severe AD. All treatments were applied twice per day. Treatment durations ranged from 28 to 42 days.

Six RCTs were vehicle-controlled, and three included active comparators. The vehicles were formulated with different emollient properties, and none of the included RCTs reported on the contents of the vehicle or the proportion of the ingredients. The contents were assumed to be

identical to the contents of the base used for the interventions.

**Outcome Assessments**

Clinical disease severity was assessed by ISGA. Three RCTs reported using a five-point ISGA scale, whereas four used a six-point scale, although scores of 0 = clear or 1 = almost clear were defined similarly across scales. Although ISGA was evaluated at various time points, the 28- to 42-day time point was the primary point of interest.

## Risk-of-Bias Assessments

Most of the RCTs were of good quality, with a low risk of bias; there were few concerns regarding the level of bias (Supplement Figure S1).

## Network Meta-Analysis Results

Six sets of network meta-analyses were performed. Because of the observed variation of vehicle effect (baseline risk), analyses adjusting for baseline risk were conducted with or without class effect. A complementary log-log (clog-log) link was used to adjust for different follow-up durations across RCTs. After fitting a variety of statistical models, the clog-log model with adjustment for baseline risk and class effects (model 1) was deemed the most appropriate because it had the lowest DIC and adjusted for baseline risk and class effects (see Table S6).

### Efficacy: ISGA 0/1 at 28–42 Days

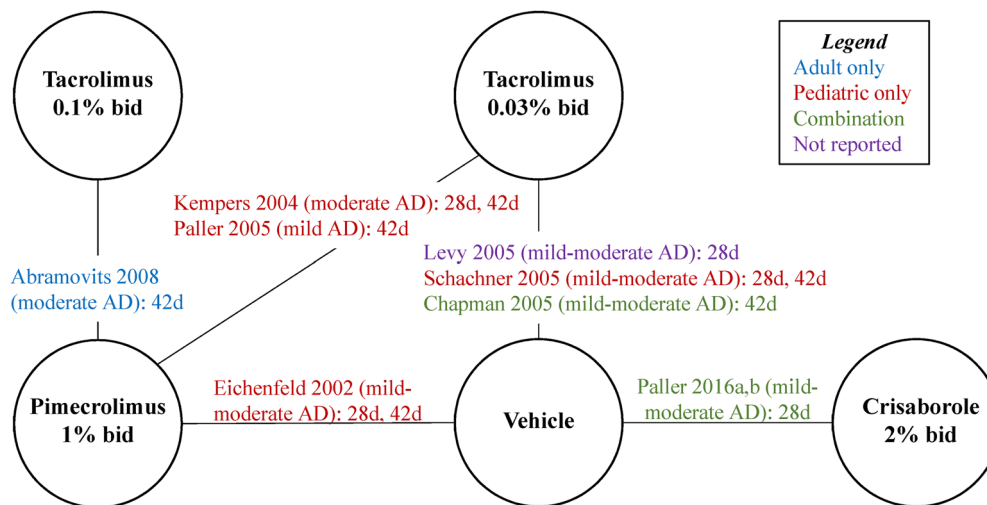
For ISGA 0/1 at 28–42 days, the clog-log model adjusting for pooled evaluation time points, baseline risk, and class effects had the lowest DIC and a residual deviance suggesting good fit (Fig. 2). As expected, the baseline risk model found strong evidence of a relationship between vehicle effect and relative treatment effect

versus vehicle (slope:  $-0.89$  [95% credible interval  $-1.26$  to  $-0.47$ ]) (Fig. 3, Table 1).

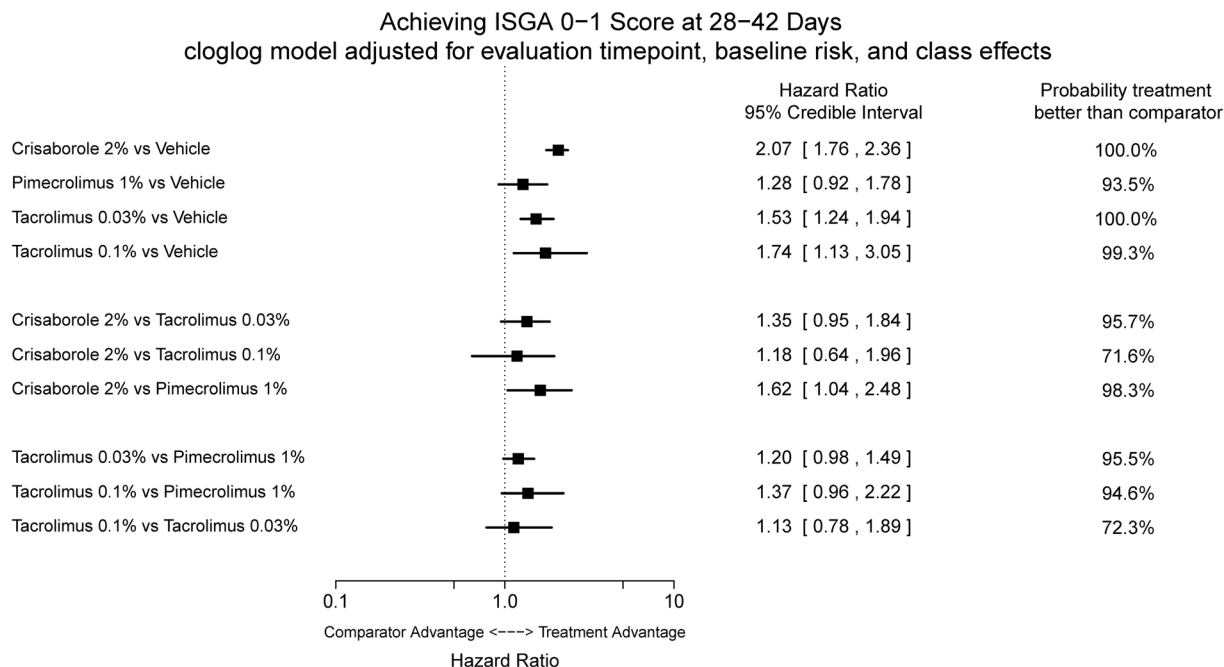
Patients on crisaborole or tacrolimus, 0.1% or 0.03%, were more likely to achieve an ISGA 0/1 at 28–42 days versus vehicle (i.e., 95% credible interval did not include 1), with the greatest point estimate observed for the crisaborole comparison (HR: 2.07; 95% credible interval 1.76–2.36; probability HR above 1 [ $p$  better]: 100.0%); there was weak evidence of a difference between pimecrolimus and vehicle (1.28; 0.92–1.78; 93.5%). Patients on crisaborole were also more likely to achieve ISGA 0/1 versus pimecrolimus (1.62; 1.04–2.48; 98.3%). There was weak evidence of a difference between crisaborole and tacrolimus, 0.03% (1.35; 0.95–1.84; 95.7%) and no evidence of a difference with tacrolimus, 0.1%.

### Safety

A network meta-analysis of safety outcomes was infeasible because of differences in reporting of safety data for comparators (e.g., different thresholds used [AEs in  $\geq 1\%$ ,  $\geq 10\%$ ]), outcome definitions (e.g., definitions of withdrawal because of AEs), and study periods between RCTs (changes in reporting of outcomes data over time; older vs. newer RCTs). Additional reasons were outcomes not reported (difficult to determine whether an outcome is not reported



**Fig. 2** Evidence network for ISGA 0/1 at 28–42 days. For studies that reported data at both 28 and 42 days, the 28-day data were used in the analyses. AD atopic dermatitis, bid twice per day, d day, ISGA Investigator's Static Global Assessment



**Fig. 3** ISGA 0/1 at 28–42 days (model 1: clog-log model adjusted for baseline risk and class effects). *ISGA* Investigator’s Static Global Assessment

because of the threshold/definitions or the outcome not occurring) and the overall sparsity of safety data reported in the included RCTs. Misalignment in the type of data reported could bias the results of any comparative quantitative analyses and might lead to under- or overestimation of results. Therefore, safety results are described qualitatively.

**Overall Adverse Events**

The rates of overall AEs ranged from 15.4% [26] to 55.6% [28]. The rates of patients reporting at least one treatment-emergent AE with crisaborole (29.3% and 29.4%) were similar to the rates experienced in the vehicle group (19.8% and 32.0%) [14]. Rates of overall AEs reported for tacrolimus, 0.03%, ranged from 15.4% [26] to 55.6% [25] across three RCTs and for tacrolimus, 0.1%, was 32.7% in one RCT [27]. These rates were 16.6% [26] to 44.0% [24] for pimecrolimus across three RCTs.

**Common Adverse Events**

Frequently reported AEs were application site burning/stinging, upper respiratory tract

infections, skin infections, and erythema (Table 1). The incidence of application site burning/stinging varied across studies and depended on the outcome definition: some studies included pain or warmth, whereas others reported only burning or stinging. Rates of application site pain AEs were 6.2% [AD-301] and 2.7% [AD-302] versus 1.2% for vehicle in each study [14], 1.9% for tacrolimus, 0.03% versus 1.8% for pimecrolimus [26], and 3.1% for tacrolimus 0.1% versus 0% for pimecrolimus [27]. Only three RCTs reported the rates of upper respiratory tract infections (2.0% [14] to 14.2% [24]). The incidence of skin infections across all RCTs was generally low (Table 1). The incidence of erythema ranged from 0% [14, 28] to 18.9% [29], but with various definitions of erythema (Table 1).

**DISCUSSION**

This systematic literature review and network meta-analysis were undertaken to evaluate the comparative effectiveness and safety of crisaborole versus other topical pharmacologic



**Table 1** Qualitative summary of safety data of included trials

Author, year (trial name)	Intervention/comparator	Adverse events, <i>n/N</i> (%) <sup>a</sup>				
		Overall	Application site	URTI	Skin infection	Erythema
Paller (AD-301) [14]	Crisaborole, 2%	TEAE: 147/502 (29.3)	Reaction: 48/502 (9.6) Pain: 31/502 (6.2)	14/502 (2.8)	Staphylococcal: 1/502 (0.2)	Application site: 1/502 (0.2)
	Vehicle	TEAE: 50/252 (19.8)	Reaction: 12/252 (4.8) Pain: 3/252 (1.2)	10/252 (4.0)	Staphylococcal: 0/252 (0)	Application site: 0/252 (0)
Paller (AD-302) [14]	Crisaborole, 2%	TEAE: 150/510 (29.4)	Reaction: 27/510 (5.3) Pain: 14/510 (2.7)	16/510 (3.1)	Staphylococcal: 1/510 (0.2)	Application site: 1/510 (0.2)
	Vehicle	TEAE: 79/247 (32.0)	Reaction: 13/247 (5.3) Pain: 3/247 (1.2)	5/247 (2.0)	Staphylococcal: 4/247 (1.6)	Application site: 0/247 (0)
Abramovits [27]	Tacrolimus, 0.1%	32/98 (32.7)	Pain: 3/98 (3.1) Burning: 19/98 (19.4)	NR	0/98 (0)	Application site: 1/98 (1.0)
	Pimecrolimus 1%	21/90 (23.3)	Pain: 0/90 (0) Burning: 12/90 (13.3)	NR	1/90 (0.1)	Application site: 2/90 (2.0)
Schachner [29]	Tacrolimus, 0.03%	NR/158 (36.7)	30/158 (19.0)	NR	NR	12/158 (7.6)
	Vehicle	NR/159 (45.3)	27/159 (17.0)	NR	NR	30/159 (18.9)
Chapman [30]	Tacrolimus, 0.03%	NR	NR	NR	NR	NR
	Vehicle	NR	NR	NR	NR	NR
Paller [26]	Tacrolimus, 0.03%	32/208 (15.4)	Pain: 4/208 (1.9) Burning: 11/208 (5.3)	NR	0/208 (0)	Application site: 2/208 (1.0)
	Pimecrolimus 1%	36/217 (16.6)	Pain: 4/217 (1.8) Burning: 20/217 (9.2)	NR	0/217 (0)	Application site: 4/217 (1.8)

**Table 1** continued

Author, year (trial name)	Intervention/comparator	Adverse events, <i>n/N (%)</i> <sup>a</sup>				
		Overall	Application site	URTI	Skin infection	Erythema
Eichenfield [24]	Pimecrolimus 1%	NR/136 (44.0)	Burning: NR (10.4)	NR/ 136 (14.2)	NR	NR
	Vehicle	NR/267 (42.6)	Burning: NR (12.5)	NR/ 267 (13.2)	NR	NR
Kempers [28]	Pimecrolimus 1%	NR	Warmth/stinging/ burning: 14/71 (20.0)	NR	Staphylococcal: 3/71 (4)	Application site: 0/71 (0)
	Tacrolimus, 0.03%	NR	Warmth/stinging/ burning: 12/70 (17.0)	NR	Staphylococcal: 0/70 (0)	Application site: NR/70 (3.0)
Levy [25]	Tacrolimus, 0.03%	NR/44 (55.6)	NR	NR	NR	NR
	Vehicle	NR/45 (48.8)	NR	NR	NR	NR

*AE* adverse event, *NR* not reported, *TEAE* treatment-emergent adverse event, *URTI* upper respiratory tract infection

<sup>a</sup> Reporting of AEs varied across studies, with some studies reporting only those AEs experienced by  $\geq 5\%$  or  $\geq 10\%$  of patients in a study arm

therapies for the treatment of mild-to-moderate AD. In the systematic literature review, no studies were identified that compared crisaborole to other active treatments. Consequently, a network meta-analysis indirectly compared treatments for which no head-to-head trials were available and synthesized available evidence across treatments. No studies of TCSs were identified that reported data on ISGA 0/1; therefore, they were not included in the network meta-analysis.

With respect to efficacy, slightly different versions of the ISGA scale were used among the RCTs. The crisaborole trials used a five-point ISGA scale as an endpoint, whereas other trials evaluated a six-point ISGA scale. Despite this, disease severity measured by baseline ISGA reported across the RCTs seemed to be comparable, with most patients having baseline ISGAs of 2–3 (mild-to-moderate). We have assumed

that the “clear” (ISGA = 0) and “almost clear” (ISGA = 1) categories are similar for both scales for analysis purposes because treatment response is defined similarly across both scales. A high response in the vehicle arm in the crisaborole trials was observed with respect to ISGA 0/1, which was greater than that seen in the vehicle arms of most RCTs that evaluated other topical therapies. This suggests that vehicle preparations in some of the RCTs do not have as many therapeutic benefits as those administered in crisaborole RCTs. Heterogeneity in patient characteristics [22], difference in the season when trials were conducted [31], and differences in the potency between creams and ointments [32] may have modified observed treatment effects. Properties of vehicle formulations may affect drug delivery and efficacy, as well as drug tolerance profiles [33]. Some vehicle excipients have a more pronounced

therapeutic effect on the skin and can improve clinical appearance and skin barrier function directly [33].

There was strong evidence that patients treated with crisaborole or tacrolimus, 0.1% or 0.03%, were more likely to achieve ISGA 0/1 at 28–42 days than those receiving vehicle. Furthermore, there was evidence that patients treated with crisaborole were more likely to achieve ISGA 0/1 at 28–42 days than those treated with pimecrolimus 1%. Although there was weak evidence of a difference between crisaborole 2% and tacrolimus 0.03%, and no evidence of a difference with tacrolimus 0.1% in model 1, all point estimates favored crisaborole.

Our findings are roughly consistent with other reported network meta-analyses on crisaborole in patients with mild-to-moderate AD; however, this may be limited given that other studies did not adjust for baseline risk (variation in efficacy rates for vehicle) [34]. The Institute for Clinical and Economic Review (ICER) report suggested that pimecrolimus was trending as superior to crisaborole [34]. However, the results of their analyses showed wide credible intervals and showed no or little evidence of any possible difference in efficacy between treatments. Although the authors of the ICER report noted there was a substantial difference in baseline risk across RCTs regarding treatment response for vehicles, they did not adjust for this in their analyses. The NICE Decision Support Unit recommends regression on baseline response as a means of adjusting for heterogeneity where appropriate [35], and, in the present case, the credible interval for the interaction term was far from zero, with a slope of  $-0.89$  and a 95% credible interval of  $-1.26$  to  $-0.47$ . In the Drug Effectiveness Review Project review, significantly more patients had treatment response with crisaborole than with vehicle [36]. The authors of this report also did not perform any adjusted analyses. As stated previously, a recent systematic literature review and network meta-analysis for PDE4is that included crisaborole and other PDE4is versus vehicle showed that topical PDE4is, particularly crisaborole, were more effective than vehicle alone [15].

Safety outcomes were not analyzed by means of a network meta-analysis in the present study because this was deemed inappropriate for a variety of reasons (e.g., difference in outcome definitions, sparsity of data). Therefore, the results for safety were only described qualitatively, and no definite conclusions regarding relative safety of crisaborole versus TCIs could be drawn. Caution should be taken in the interpretation of naive comparisons because no formal comparative (indirect) assessments were made.

The strengths of our study include various key aspects relative to the innovative application of meta-analysis methodologies to address the need for comparative efficacy evidence. The systematic literature review was performed in accordance with published guidelines, and the network meta-analysis was based on well-established Bayesian methodology [20, 37, 38]. Our systematic review and network meta-analysis was rigorous, used sophisticated statistical models, and reached conclusions that have not been previously documented. Heterogeneity was addressed, where possible, to fulfill the homogeneity assumption necessary for the network meta-analysis. A comprehensive feasibility assessment was conducted a priori, including an evaluation of the clinical heterogeneity between trials that showed that studies were similar for many of the characteristics of interest. Baseline risk regression was performed to adjust for differences in vehicle response and heterogeneity in treatment effects across trials.

There are several limitations to this study. First, the interval for the primary time point of interest was wide at 28–42 days. Because efficacy for interventions may change with prolonged use, this is a potential source of heterogeneity and may have impacted the results for this outcome (i.e., ISGA 0/1 at 28–42 days). To control for this variability in follow-up time, a clog-log model was applied for ISGA 0/1 at the 28- to 42-day time point.

A second limitation refers to the efficacy data being evaluated, given possible confounding factors and the issue that data for some other efficacy outcomes also important in AD were not available. The efficacy difference may not be generalizable to some real clinic settings, as

there may be other confounding factors associated with the use and benefits of active treatment in real clinic settings (e.g., access issues). It was also not possible to fully explore all potential confounders by means of subgroup analyses. Further adjustment for differences in baseline characteristics could not be explored using meta-regression techniques because of the limited number of studies available for comparators. Also, some important efficacy outcomes could not be evaluated because of data limitations (e.g., pruritus reduction, quality of life benefit).

A third limitation was that safety outcomes could only be described qualitatively. Network meta-analysis for safety was inappropriate because of sparse data across studies, including differences in outcome definitions used, in reporting of data for comparators, issues with outcomes not reported, and differences in study period.

There are no head-to-head trials comparing crisaborole versus other active treatments. We could only indirectly compare treatments using network meta-analysis. Results should be interpreted with caution and cannot replace a direct head-to-head evaluation.

## CONCLUSION

This network meta-analysis showed that crisaborole was superior to vehicle and pimecrolimus and comparable to tacrolimus, 0.1% or 0.03% in achieving ISGA 0/1 at 28–42 days in patients aged  $\geq 2$  years with mild-to-moderate AD. In the crisaborole pivotal studies (AD-301/AD-302), crisaborole was shown to be well tolerated, with low rates of treatment-related AEs. More research is needed to establish the comparative efficacy of crisaborole with respect to other key clinical efficacy outcomes, including other severity scales (e.g., Eczema Area and Severity Index [EASI], the SCORing Atopic Dermatitis [SCORAD]) and assessments of pruritus severity, in addition to other patient-relevant outcomes (e.g., QOL and functional status).

## ACKNOWLEDGEMENTS

We thank Dr. Jacob Thyssen for his expert clinical interpretations of study results and critical review of the manuscript. We also thank Alison Chapman and Dr. William Romero for their expert review of the manuscript with respect to both technical methods and clinical aspects. We also thank Caroline Cole, Janet Dooley, and Lauren Randall of Evidera for their very valuable contributions in providing medical writing, editing, and formatting support for this manuscript.

**Funding.** Funding for this study, this manuscript (including medical writing, editing, and manuscript formatting support), and the Rapid Service Fee was provided by Pfizer Inc. (New York, NY, USA).

**Authorship.** All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this manuscript, take responsibility for the integrity of the work as a whole, and have provided final approval of the version to be published

**Medical Writing, Editorial, and Other Assistance.** Medical writing and editorial and manuscript formatting support was provided by Caroline Cole, Janet Dooley, and Lauren Randall of Evidera and was funded by Pfizer.

**Prior Presentation.** Some of the data in this manuscript have appeared in a poster presented at the International Society for Pharmacoeconomics and Outcomes Research 22nd Annual European Congress; 2–6 November 2019; Copenhagen, Denmark. Fahrback K, Tarpey J, Bergrath E, et al. A systematic literature review and network meta-analysis (NMA) of crisaborole 2% ointment for the treatment of mild-to-moderate atopic dermatitis (AD). Value Health. 2019;22(Suppl 3):S597–S598. Abstract PDG6.

**Disclosures.** Maureen P. Neary, Amy Cha, Robert Gerber, and Joseph C. Cappelleri are employed by Pfizer Inc. and own Pfizer stock. Howard Thom has received personal consulting fees from Pfizer Inc., Novartis Pharma AG,

Roche Holding AG, Bayer AG, and Janssen Pharmaceuticals. Kyle Fahrback, Jialu Tarpey, Evelien Bergrath, and Rachel Hughes are employed by Evidera, which provides consulting and other research services to pharmaceutical, medical device, and related organizations. In their salaried positions, they work with a variety of companies and organizations and are precluded from receiving payment or honoraria directly from these organizations for services rendered. Evidera received funding from Pfizer Inc. to conduct this study and develop the manuscript.

**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. The review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.

**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. Bieber T. Atopic dermatitis. *N Engl J Med.* 2008;358(14):1483–94.
2. Bieber T. Atopic dermatitis. *Ann Dermatol.* 2010;22(2):125–37.
3. Schneider L, Tilles S, Lio P, et al. Atopic dermatitis: a practice parameter update 2012. *J Allergy Clin Immunol.* 2013;131(2):295–299.e27.
4. Flohr C, Mann J. New insights into the epidemiology of childhood atopic dermatitis. *Allergy.* 2014;69(1):3–16.
5. Silverberg JI, Simpson EL. Associations of childhood eczema severity: a US population-based study. *Dermatitis.* 2014;25(3):107–14.
6. Asthma and Allergy Foundation of America (AAFA). Atopic dermatitis in America Study Overview. 2018. <https://www.aafa.org/media/2209/Atopic-Dermatitis-in-America-Study-Overview.pdf>. Accessed 19 July 2019.
7. Chiesa Fuxench ZC, Block JK, Boguniewicz M, et al. Atopic dermatitis in America Study: a cross-sectional study examining the prevalence and disease burden of atopic dermatitis in the US adult population. *J Invest Dermatol.* 2019;139(3):583–90.
8. Eichenfield LF, Tom WL, Berger TG, et al. Guidelines of care for the management of atopic dermatitis: section 2. Management and treatment of atopic dermatitis with topical therapies. *J Am Acad Dermatol.* 2014;71(1):116–32.
9. Andersen YMF, Egeberg A, Ban L, et al. Association between topical corticosteroid use and type 2 diabetes in two European population-based adult cohorts. *Diabetes Care.* 2019;42(6):1095–103.
10. Charman CR, Morris AD, Williams HC. Topical corticosteroid phobia in patients with atopic eczema. *Br J Dermatol.* 2000;142(5):931–6.
11. Walling HW, Swick BL. Update on the management of chronic eczema: new approaches and emerging treatment options. *Clin Cosmet Investig Dermatol.* 2010;3:99–117.
12. Hanifin JM, Lloyd R, Okubo K, Guerin LL, Fancher L, Chan SC. Relationship between increased cyclic AMP-phosphodiesterase activity and abnormal adenylyl cyclase regulation in leukocytes from patients with atopic dermatitis. *J Invest Dermatol.* 1992;98(6 Suppl):100S–5S.
13. Gittler JK, Shemer A, Suarez-Farinas M, et al. Progressive activation of T(H)2/T(H)22 cytokines and

- selective epidermal proteins characterizes acute and chronic atopic dermatitis. *J Allergy Clin Immunol.* 2012;130(6):1344–54.
14. Paller AS, Tom WL, Lebwohl MG, et al. Efficacy and safety of crisaborole ointment, a novel, nonsteroidal phosphodiesterase 4 (PDE4) inhibitor for the topical treatment of atopic dermatitis (AD) in children and adults. *J Am Acad Dermatol.* 2016;75(3):494–503.e6.
  15. Yang H, Wang J, Zhang X, et al. Application of topical phosphodiesterase 4 inhibitors in mild to moderate atopic dermatitis: a systematic review and meta-analysis. *JAMA Dermatol.* 2019;155(5):585–93.
  16. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med.* 2015;162(11):777–84.
  17. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med.* 2009;6(7):e1000100.
  18. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ.* 2011;343:d5928.
  19. National Institute for Health and Care Excellence (NICE). Single technology appraisal: user guide for company evidence submission template. 2015. <https://www.nice.org.uk/process/pmg24/chapter/instructions-for-companies>. Last updated April 2017. Accessed March 26, 2020.
  20. Dias S, Ades AE, Welton NJ, Jansen JP, Sutton AJ. Network meta-analysis for decision-making. Chichester: Wiley; 2018.
  21. Lunn D, Jackson C, Best N, Spiegelhalter D, Thomas A. The BUGS book: a practical introduction to Bayesian analysis. Boca Raton: CRC Press; 2012.
  22. Achana FA, Cooper NJ, Dias S, et al. Extending methods for investigating the relationship between treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis. *Stat Med.* 2013;32(5):752–71.
  23. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Series B Stat Methodol.* 2002;64(4):583–639.
  24. Eichenfield LF, Lucky AW, Boguniewicz M, et al. Safety and efficacy of pimecrolimus (ASM 981) cream 1% in the treatment of mild and moderate atopic dermatitis in children and adolescents. *J Am Acad Dermatol.* 2002;46(4):495–504.
  25. Levy AL, Sheehan M, Roberts R. Tacrolimus cream 0.03% is safe and effective in the treatment of mild to moderate atopic dermatitis in adults. *J Allergy Clin Immunol.* 2005;115(2):S103.
  26. Paller AS, Lebwohl M, Fleischer AB Jr, US/Canada Tacrolimus Ointment Study Group, et al. Tacrolimus ointment is more effective than pimecrolimus cream with a similar safety profile in the treatment of atopic dermatitis: results from 3 randomized, comparative studies. *J Am Acad Dermatol.* 2005;52(5):810–22.
  27. Abramovits W, Fleischer AB Jr, Jaracz E, Breneman D. Adult patients with moderate atopic dermatitis: tacrolimus ointment versus pimecrolimus cream. *J Drugs Dermatol.* 2008;7(12):1153–8.
  28. Kempers S, Boguniewicz M, Carter E, et al. A randomized investigator-blinded study comparing pimecrolimus cream 1% with tacrolimus ointment 0.03% in the treatment of pediatric patients with moderate atopic dermatitis. *J Am Acad Dermatol.* 2004;51(4):515–25.
  29. Schachner LA, Lamerson C, Sheehan MP, US Tacrolimus Ointment Study Group, et al. Tacrolimus ointment 0.03% is safe and effective for the treatment of mild to moderate atopic dermatitis in pediatric patients: results from a randomized, double-blind, vehicle-controlled study. *Pediatrics.* 2005;116(3):e334–e342.
  30. Chapman MS, Schachner LA, Breneman D, US Tacrolimus Ointment Study Group, et al. Tacrolimus ointment 0.03% shows efficacy and safety in pediatric and adult patients with mild to moderate atopic dermatitis. *J Am Acad Dermatol.* 2005;53(2 Suppl 2):S177–S185185.
  31. Engebretsen KA, Kezic S, Riethmuller C, et al. Changes in filaggrin degradation products and corneocyte surface texture by season. *Br J Dermatol.* 2018;178(5):1143–50.
  32. Ference JD, Last AR. Choosing topical corticosteroids. *Am Fam Physician.* 2009;79(2):135–40.
  33. Abramovits W, Boguniewicz M, Paller AS, et al. The economics of topical immunomodulators for the treatment of atopic dermatitis. *Pharmacoeconomics.* 2005;23(6):543–66.
  34. Institute for Clinical and Economic Review (ICER). Dupilumab and Crisaborole for Atopic Dermatitis: Effectiveness and Value: Final Evidence Report and Meeting Summary. 2017. <https://icer-review.org/>

- 
- [material/atopic-dermatitis-final-report/](#). Published June 8, 2017. Last Accessed July 19, 2019.
35. Dias S, Sutton AJ, Welton NJ, Ades AE. NICE DSU Technical support document 3: heterogeneity: subgroups, meta-regression, bias and bias-adjustment. Report by the decision support unit. 2011. <https://www.nicedsu.org.uk>. Published September 2011. Last updated April 2012. Accessed February 10, 2020.
  36. Oregon Health and Science University, Drug Effectiveness Review Project. Drugs to Treat Atopic Dermatitis: Final Report. 2017.
  37. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res.* 2001;10(4):277–303.
  38. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med.* 2004;23(20):3105–24.