CADTH

CADTH COMMON DRUG REVIEW

Clinical Review Report

CRISABOROLE Ointment, 2% (EUCRISA)

(Pfizer Canada Inc.)

Indication: For topical treatment of mild to moderate atopic dermatitis in patients 2 years of age and older.

Service Line:	CADTH Common Drug Review
Version:	Final with Redactions
Publication Date:	April 2019
Report Length:	72 Pages

CADTH

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

CADTH

Table of Contents

Tables

Figures

Abbreviations

AD	atopic dermatitis
AE	adverse event
CDLQI	Children's Dermatology Life Quality Index
CI	confidence interval
CrI	credible interval
CSPA	Canadian Skin Patient Alliance
DFI	Dermatitis Family Impact questionnaire
DLQI	Dermatology Life Quality Index
ESC	Eczema Society of Canada
GQ	Global Question
GRCQ	Global Rating of Change Questionnaire
HRQoL	health-related quality of life
IDC	indirect treatment comparison
ISGA	Investigator's Static Global Assessment
ITT	intention-to-treat population
MCMC	Markov Chain Monte Carlo
MCID	minimal clinically important difference
NMA	network meta-analysis
PDE4	phosphodiesterase type 4
QoL	quality of life
RCT	randomized controlled trial
SAE	serious adverse event
SD	standard deviation
SOC	standard of care
TCI	topical calcineurin inhibitor
TCS	topical corticosteroids
WDAE	withdrawal due to adverse event

Drug	Crisaborole (Eucrisa)
Indication	Crisaborole (Eucrisa ointment, 2%) is indicated for topical treatment of mild to moderate atopic dermatitis in patients 2 years of age and older
Reimbursement Request	For the treatment of mild to moderate atopic dermatitis in patients 2 years of age and older who have failed or are intolerant to a topical corticosteroid treatment
Dosage Form(s)	Ointment, 2% for topical use
NOC Date	June 7, 2018
Manufacturer	Pfizer Canada Inc.

Executive Summary

Introduction

Atopic dermatitis (AD) is a chronic, relapsing, and inflammatory skin condition, characterized by eczematous lesions, pruritus, and dry skin. Pruritus of the skin causes frequent scratching and may result in lichenification (thickening of the skin) and secondary skin infections. The symptoms of AD wax and wane and disease severity can range from mild to severe disease. AD begins in early childhood with the majority of cases beginning before the age of five years.[1,2] Although childhood symptoms resolve by adolescence, some patients' AD symptoms will persist or develop in adulthood.[3,4] The Canadian Dermatology Association reported that the lifetime prevalence of AD is up to 17%, and there is evidence to suggest that the prevalence has increased over the past 30 years.

The goal of AD management is to prevent and manage flare-ups, which are recurrent episodes of worsening of symptoms that require an escalation of treatment.[3] Although there is no cure for AD, there are several therapeutic options available to patients. The majority of patients treat AD using general skin care methods and topical anti-inflammatory therapies. However, if these practices fail to improve AD symptoms, patients may use off-label systemic immune-modulating agents or other therapies, such as phototherapy. The most commonly pharmaceutical topical therapies for patients with AD include topical corticosteroids (TCS) and topical calcineurin inhibitors (TCI). TCS are anti-inflammatory agents that act to control flare-ups and they are considered a first-line therapy for patients with AD. Side effects associated with long-term use of TCS include striae (stretch marks), petechiae (small red or purple spots), telangiectasia (small, dilated blood vessels on the surface of the skin), skin thinning, atrophy, and acne.[4] On the other hand, TCI are steroid-free, anti-inflammatory, immunosuppressant agents. In Canada, TCIs are used in the second-line setting for patients who exhibit steroid phobia or where the use of steroids is not advisable. The most common adverse event associated with TCI therapy is application site–specific burning and irritation.

Crisaborole is a low-molecular-weight benzoxaborole, nonsteroidal, topical ointment. Crisaborole inhibits phosphodiesterase type 4 (PDE4), which regulates inflammatory cytokine production. It is applied in a thin layer to the affected area, twice daily.

The current CADTH Common Drug Review (CDR) submission for crisaborole is for the treatment of patients two years of age and older with mild-to-moderate AD.

Results and Interpretation

Included Studies

Two multi-centre, manufacturer-sponsored, double-blind randomized controlled trials (RCTs) met the inclusion criteria for this CDR review. Studies AD-301 (N = 763) and AD-302 (N = 764) were identically designed trials and enrolled patients two years of age and older (majority under 18 years of age) with mild-to-moderate AD (Investigator's Static Global Assessment [ISGA] scoring) comparing crisaborole in a 2:1 ratio to vehicle over a 28-day treatment course. The primary outcome was the proportion of patients with success by ISGA at day 29, while secondary outcomes included the proportion of patients with ISGA of clear or almost clear at day 29, and the time to success in ISGA. The ISGA is a 5-point scale that provides a global clinical assessment of AD severity based on an ordinal scale, scored by an investigator or physician ranging from 0 to 4. A score of 0 corresponds to a grade of clear; 1 is almost clear; 2 is mild; 3 is moderate; and 4 is severe AD. A decrease in score relates to an improvement in signs and symptoms. No minimal clinically important difference (MCID) is available for ISGA in patients with AD.

Health-related quality of life (HRQoL) was not statistically assessed in either included study, therefore, no conclusions can be drawn about the impact of crisaborole on quality of life, an important consideration for patients with AD. There were numerically more withdrawals in the vehicle group than with crisaborole, and this difference appears to have been largely accounted for by lack of efficacy. The included trials lacked an active comparator, therefore the relative efficacy of crisaborole to TCS or TCI is unknown. The included studies had a relatively short duration of follow-up and thus long-term efficacy and safety of crisaborole is unknown.

Efficacy

The primary outcome in both trials was the proportion of patients achieving success in ISGA at day 29. Success was defined as ISGA of clear or almost clear with at least a 2-grade improvement from baseline/day 1. For both trials, patients treated with crisaborole were significantly more likely to have a success as compared with those treated with vehicle (AD-301 = 32.8% versus 25.4%, *P* value: 0.038 and AD-302= 31.4% versus 18.0%; *P* value < 0.001). [redacted] Subgroup analyses were presented based on age (less than 18 years versus greater than and equal to 18 years of age), however, no interaction test results were reported and no formal comparisons were conducted between treatment groups. There was no clear indication that patient age impacts response to crisaborole. Additionally, in response to a request from CDR, a post hoc pooled subgroup analysis was presented with responses broken down by baseline AD severity (ISGA, mild versus moderate). In this case the treatment effect was not statistically significant in patients with mild AD (success in 24.9% of crisaborole patients versus 21.1% of vehicle patients; mean difference between groups of 3.6 [95% confidence interval (CI), –3.9 to 11.2; *P* = 0.35]), while it was statistically significant in patients with moderate disease (36.7% versus 22.3% respectively; mean difference between groups of 14.4% [95% CI, 8.0 to 20.8; *P* < 0.0001]).

HRQoL was assessed using the Dermatology Life Quality Index (DLQI) in both studies, although no statistical tests were performed between comparison groups. In AD-301, the mean standard deviation (SD) decrease (improvement) from baseline to day 29 was –5.5 (5.5) for crisaborole and –3.6 (4.6) for vehicle. In AD-302 the mean (SD) decrease from baseline to day 29 was –5.0 (5.5) for crisaborole and –3.4 (5.8) with vehicle. The MCID for a change from baseline is 3.3, thus clinically significant improvement from baseline was seen in both the crisaborole and vehicle groups.

With respect to the children's DLQI, the mean (SD) reduction from baseline to day 29 in AD-301 was –5.2 (5.6) with crisaborole and –3.1 (5.9) with vehicle, and in AD-302 was –4.0 (4.9) with crisaborole and –2.9 (5.0) with vehicle.

The mean (SD) reduction (improvement) from baseline to day 29 in Dermatitis Family Impact questionnaire (DFI) in AD-301 was –3.9 (5.7) with crisaborole and –2.7 (5.6) with vehicle, and in AD-302 it was –3.6 (5.2) with crisaborole and –2.8 (4.8) with vehicle.

The median time to improvement in pruritus was an exploratory end point in both included studies. The median time to improvement in pruritus was 1.32 days with crisaborole and 1.87 days with vehicle ($P < 0.001$) in AD-301, and 1.41 days with crisaborole and 1.54 days with vehicle ($P = 0.425$) in AD-302.

Harms

There were no deaths in either study.

Adverse events (AEs) were reported in 29% of patients in the crisaborole group and 20% of patients in the vehicle group in AD-301, and in 29% of crisaborole and 30% of vehicle-treated patients in AD-302. Application site pain was the most common AE in AD-301, occurring in 6.2% of crisaborole-treated versus 1.2% of vehicle-treated patients after 29 days. In AD-302, application site pain occurred in 2.7% of crisaborole-treated and 1.2% of vehicle-treated patients after 29 days.

There were few serious adverse events (SAEs) through the 29-day treatment period of either study. In AD-301, 1.0% of crisaborole-treated and 0.4% of vehicle-treated patients had an SAE, while in AD-302 0.6% of crisaborole-treated patients experienced an SAE, and none in the vehicle group.

In AD-301, withdrawal due to adverse event (WDAE) occurred in 1.4% of crisaborole-treated patients and in 0.8% of vehicle-treated patients and in AD-302 occurred in 1.0% of crisaborole patients and 1.6% of vehicle-treated patients. The only WDAE that occurred in more than one patient in either study was application site pain and application site urticaria, each occurring in two patients in the crisaborole group in AD-301.

With respect to notable harms, one patient had a hypersensitivity reaction, reported as a treatment-emergent AE, in the crisaborole group in AD-302. Otherwise there were no other hypersensitivity reactions reported.

Potential Place in Therapy[1]

Based on current therapies and standard of care for AD, an unmet need would be effective, affordable, and safe therapies for patients suffering with the disease. One challenge would be steroid phobia. While this would best be addressed through proper patient education about appropriate use of topical steroids, crisaborole would add to the armamentarium of treatment options for these patients. Another challenge would be topical treatment options for unresponsive and/or severe AD patients. In this respect, it is unknown whether or not crisaborole would meet this need. With respect to affordability, the projected cost of approximately $2.00 per gram of medication makes it comparable with TCI and does not offer an advantage as compared with TCS.

There are some fears about the black box labelling of TCI. While this labelling may generate more fear in the patient population about using TCI, the real-world safety is likely not impacted by switching to crisaborole. Nonetheless, there may be a perceived safety by patients in using crisaborole.

Patients with mild-to-moderate AD who are unwilling to use topical steroids (e.g., for reasons such as steroid phobia) would be good candidates to receive topical crisaborole. No special diagnostic tests would need to be run.

Conclusions

Two identically designed multi-centre, double-blind RCTs, AD-301 and AD-302, both entirely based in the US, met the inclusion criteria for this systematic review. Both studies randomized patients with AD scored as mild-to-moderate using ISGA, in a 2:1 ratio to either crisaborole or vehicle over a treatment course of 28 days. A larger percentage of crisaborole-treated patients versus vehicle achieved the primary outcome of treatment success according to the ISGA at day 29, and this difference was statistically significant in both studies. No conclusions can be drawn regarding HRQoL, and this is an important limitation given the impact of AD on this outcome. There was some indication of a numerically higher risk of application site pain with crisaborole, although there was no increased risk of WDAE. The lack of an active comparator in the two trials was a limitation of this review. [redacted] however, there were several limitations with the analysis, and no data were available to compare crisaborole to TCS.

[1] This information is based on information provided in draft form by the clinical expert consulted by CDR reviewers for the purpose of this review.

Table 1: Summary of Results

Efficacy Outcomes	AD-301		AD-302	
Primary End Points	**Crisaborole N = 503**	**Vehicle N = 256**	**Crisaborole N = 513**	**Vehicle N = 250**
Proportion of Patients Achieving Success in ISGA at Day 29,[a] n (%)				
Baseline ISGA				
Mild (ISGA score 2)	196 (39.0%)	93 (36.3%)	197 (38.4%)	100 (40.0%)
Moderate (ISGA score 3)	307 (61.0%)	163 (63.7%)	316 (61.6%)	150 (60.0%)
Success[a]	165 (32.8%)	65 (25.4%)	161 (31.4%)	45 (18.0%)
Failure	338 (67.2%)	191 (74.6%)	352 (68.6%)	205 (82.0%)
P Value[b]	*P* = 0.038		*P* < 0.001	
Secondary Outcomes				
ISGA of Clear or Almost Clear at Day 29				
Success[a]	260 (51.7%)	104 (40.6%)	249 (48.5%)	74 (29.7%)
Failure	243 (48.3%)	152 (59.4%)	264 (51.5%)	176 (70.3%)
P value	*P* = 0.005		*P* < 0.001	
Time to Success in ISGA[a] (days)				
N	503	256	513	250
Median	NA	NA	NA	NA
P value	*P* < 0.001		*P* < 0.001	
Exploratory End Points				
Time to Improvement in Pruritus (days)				
N	428	210	439	211
Median	1.32	1.87	1.41	1.54
P value	*P* < 0.001		*P* = 0.425	
DLQI				
Baseline N	95	52	97	40
Baseline mean (SD)	9.6 (6.37)	9.5 (6.52)	9.7 (6.24)	9.1 (6.67)
Mean (SD) change from baseline to day 29	–5.5 (5.45) N = 87	–3.6 (4.60) N = 44	–5.0 (5.49) N = 93	–3.4 (4.75) N = 38
CDLQI				
Baseline N	393	199	404	204
Baseline mean (SD)	9.7 (6.19)	9.1 (6.54)	9.0 (5.77)	8.9 (5.48)
Mean (SD) change from baseline to day 29	–5.2 (5.63) N = 374	–3.1 (5.90) N = 175	–4.0 (4.92) N = 376	–2.9 (5.01) N = 180
DFI				
Baseline N	431	214	431	217
Baseline mean (SD)	8.5 (6.63)	7.5 (6.66)	7.7 (6.57)	8.0 (5.65)
Mean (SD) change from baseline to day 29	–3.9 (5.68) N = 407	–2.7 (5.61) N = 187	–3.6 (5.18) N = 404	–2.8 (4.75) N = 190
Treatable % BSA				
Baseline mean (SD)	18.8 (18.55) N = 503	18.6 (18.87) N = 256	17.9 (17.49) N = 513	17.7 (15.61) N = 250
Mean (SD) change from baseline to day 29	–8.2 (12.83)	–5.8 (12.79)	–6.7 (12.22)	–3.1 (11.07)

Efficacy Outcomes	AD-301		AD-302	
	N = 477	N = 228	N = 486	N = 224
Harms				
Subjects with > 0 AEs, N (%)	147 (29.3)	50 (19.8)	150 (29.4)	79 (32.0)
Subjects with > 0 SAEs, N (%)	5 (1.0)	1 (0.4)	3 (0.6)	0
WDAEs, N (%)	7 (1.4)	2 (0.8)	5 (1.0)	4 (1.6)

AE = adverse event; BSA = body surface area; CLDQI = Children's Dermatology Life Quality Index; DFI = Dermatitis Family Impact Questionnaire; DLQI = Dermatology Life Quality Index; ISGA = Investigator's Static Global Assessment; NA = not available; SAE = serious adverse event; SD = standard deviation.

Note: Improvement in pruritus defined as achieving none (0) or mild (1) with at least a 1-grade improvement from baseline. Medians computed using Kaplan–Meier methods.

[a] Success in ISGA for the primary outcome was defined as ISGA of clear or almost clear with at least a 2-grade improvement from baseline/day 1. This differed from the secondary end point, which defined success as simply being clear or almost clear.

[b] The *P* value from a logistic regression (with Firth option) test with factors of treatment group and analysis centre. The adjusted estimates for AD-301 and AD-302 from logistic regression were 29.1% and 22.0% and 26.5% and 14.2% for the crisaborole and vehicle groups, respectively. Values were adjusted for multiple imputation.

Source: Clinical study reports for AD-301[5] and AD-302.[6]

Introduction

Disease Prevalence and Incidence

Atopic dermatitis (AD) is a chronic, relapsing, and inflammatory skin condition. The precise mechanisms behind AD have remained elusive over the years; however, there is clearly a breakdown in the skin's barrier function and an immunologic component.[2] AD is characterized by eczematous lesions, pruritus, and dry skin. AD lesions appear as fluid-filled vesicles that ooze, crack, and crust. Pruritus of the skin causes frequent scratching and may result in lichenification (thickening of the skin) and secondary skin infections. The symptoms of AD wax and wane and disease severity can range from mild to severe disease. Additionally, the clinical presentation of AD differs depending on age. In children (two years to puberty), AD lesions appear on the flexural surfaces of extremities (e.g., folds or bends at the elbow or knee), neck, wrist, and ankles, while they may present on the flexural surfaces of the extremities, hands, and feet in adolescents and adults.[4]

AD begins in early childhood with the majority of cases beginning before the age of five years.[1,2] Although childhood symptoms resolve by adolescence, some patients' AD symptoms will persist or develop in adulthood.[3,4] The Canadian Dermatology Association reported that the lifetime prevalence of AD is up to 17%, and there is evidence to suggest that the prevalence has increased over the past 30 years.[1,3,4] AD has a negative impact on quality of life for adults, children, and their caregivers. For instance, the most common negative experiences for adults and children with AD are sleeping disturbances, anxiety, and avoidance of social activities.[7,8] Some children have reported that their condition has had a detrimental impact on school attendance.[8] Overall, patient experiences describe a physically and mentally exhausting condition that can result in anxiety, depression, and decrease in quality of life.

The goal of AD management is to prevent and manage flare-ups, which are recurrent episodes of worsening of symptoms that require an escalation of treatment.[3] Although there is no cure for AD, there are several therapeutic options available to patients. The majority of patients treat AD using general skin care methods and topical anti-inflammatory therapies. However, if these practices fail to improve AD symptoms, patients may use off-label systemic immune-modulating agents or other therapies, such as phototherapy.

Standards of Therapy

General Skin Care

It is recommended that patients with AD develop good skin care practices. First, patients should avoid common irritants, such as high temperatures, dust, smoke, grass, and chlorine or solvents.[1,3] Patients should also wash clothes with mild detergents with no fabric softener and double-rinse clothes. Secondly, patients with AD are encouraged to practice daily dry skin management, which includes bathing in lukewarm water with mild cleansing agents and regular moisturizer application.[1,3,4,9]

Topical Therapies

The most commonly recommended topical therapies for patients with AD are moisturizers. Routine moisturizer use provides a barrier for the skin from common allergens and irritants, as well as helping to soften, reduce itching, and minimize cracking, fissuring, and

lichenification of the skin.[3,9] Examples of moisturizers include a combination of emollients, humectants, and occlusive agents. Emollients (e.g., glycol, glyceryl stearate, soy sterols) act to smooth out the surface of the skin by filling in space with droplets of oils. Humectants (e.g., glycerol, lactic acid, urea) increase the water-holding capacity of skin; however, they are not recommended for children because they sting open skin. Finally, occlusive agents (e.g., petrolatum, dimethicone, mineral oil) add a layer of oil on top of the skin, which helps to reduce water loss, and thereby, increases the moisture of the skin. The type of moisturizer recommended depends on the area of the body and the dryness of the skin.[3,9]

The most commonly used pharmaceutical topical therapies for patients with AD include topical corticosteroids (TCS) and topical calcineurin inhibitors (TCI). TCS are anti-inflammatory agents that act to control flare-ups. They are considered a first-line therapy for patients with AD.[4] There is a variety of different TCS, which range in potency (e.g., low to very potent) and preparation (e.g., ointment and cream). In Canada, hydrocortisone 1% (low potency) is the most commonly prescribed type of TCS for the face.[3] For the body, triamcinolone or betamethasone valerate (moderate potency) are most commonly prescribed.[3] TCS are applied directly to the area of affected skin prior to the use of emollients, and treatment response is typically seen within 10 to 14 days. Short-course, mid- or higher-potency TCS are recommended to reduce the symptoms of acute AD flares in adults and children, while lower-potency TCS are recommended for long-term management.[9] Management of these flares can continue for weeks while maintenance therapy, which often requires application only on an intermittent basis (e.g., twice weekly), is continued perpetually. Side effects associated with long-term use of TCS include striae (stretch marks), petechiae (small red or purple spots), telangiectasia (small, dilated blood vessels on the surface of the skin), skin thinning, atrophy, and acne.[4] On the other hand, TCI are steroid-free, anti-inflammatory, immunosuppressant agents. In Canada, TCI are used in the second-line setting for patients who exhibit steroid phobia or where the use of steroids is not advisable.[3] Two TCI are currently available in Canada, pimecrolimus (1%) and tacrolimus (0.03% and 0.1%). Pimecrolimus 1% cream is effective in controlling pruritus and it can be used for short-term and intermittent long-term therapy for patients with mild-to-moderate AD.[3] TCI have been used for up to a year intermittently without significant adverse effects.[9] Topical tacrolimus is an ointment that demonstrates rapid and sustained AD symptom control and it can be used for short-term and intermittent long-term therapy of moderate-to-severe AD.[3,10] Topical tacrolimus 0.03% and 0.1% are indicated for adults and only 0.03% is indicated for children aged 2 to 15 years.[9] The most common adverse event (AE) associated with TCI therapy is application site–specific burning and irritation.[3,4]

Systemic Therapies

Systemic therapies for patients with AD include antimicrobial, antihistamine, or systemic immunomodulating agents.[11,12] Systemic antibiotic treatment can be used to counter widespread secondary bacterial infection. Many patients encounter infection with *Staphylococcus aureus* and this may cause new inflammation and exacerbate AD symptoms. The choice of systemic antibiotic agent depends upon the skin culture and sensitivity profile. Sedating antihistamines have been used in cases where patients are not achieving adequate sleep due to itching.[1,10] Systemic immunomodulating agents are recommended for patients with AD who have failed or are intolerant to TCS and/or TCI treatment. These agents include: cyclosporine A, azathioprine, methotrexate, and mycophenolate mofetil.[11-13] These agents are off-label in Canada and it has been

recommended that patients receive the lowest dose for the shortest duration to avoid adverse effects.[11,12]

Other Therapies

Phototherapy is a second-line therapy that is recommended for patients who failed to respond to first-line therapies, such as emollients, TCS, and/or TCI.[11,12] Other emerging therapies, such as dupilumab, a fully human monoclonal antibody that blocks interleukin-4 and interleukin-13, have been approved for use in Canada. Dupilumab can be used in combination with TCS for adult patients with moderate-to-severe AD that is not adequately controlled by topical therapies or when these therapies are inadvisable.

Drug

Crisaborole is a low-molecular-weight benzoxaborole, nonsteroidal, topical ointment. Crisaborole inhibits phosphodiesterase type 4 (PDE4), which regulates inflammatory cytokine production. It is applied in a thin layer to the affected area, twice daily.

The current CADTH Common Drug Review (CDR) submission for crisaborole is for the treatment of patients two years of age and older with mild-to-moderate AD.

Table 2: Key Characteristics of Crisaborole, Topical Calcineurin Inhibitors, and Corticosteroids

	Crisaborole	Tacrolimus	Pimecrolimus	Topical Corticosteroids
Mechanism of action	PDE4 inhibitor	Calcineurin inhibitor	Calcineurin inhibitor	Via multiple mechanisms, acts as an anti-inflammatory and immune suppressant
Indication[a]	Topical treatment of mild-to-moderate AD in patients 2 years of age and older.	Second-line therapy for short- and long-term intermittent treatment of moderate-to-severe AD in non-immunocompromised patients, in whom the use of conventional therapies are deemed inadvisable because of potential risks, or who are not adequately responsive to or intolerant of conventional therapies.	Second-line therapy for short-term and intermittent long-term therapy of mild-to-moderate AD in non-immunocompromised patients 2 years of age and older, in whom the use of alternative, conventional therapies is deemed inadvisable because of potential risks, or in the treatment of patients who are not adequately responsive to or intolerant of alternative, conventional therapies.	Symptomatic relief of acute and chronic skin eruptions, where anti-inflammatory, anti-allergenic, and antipruritic activity is required. Clinical evidence of effectiveness is available for AD, as well as other indications such as psoriasis, vitiligo, phimosis, acute radiation dermatitis, and lichen sclerosus; and there is limited clinical evidence for use in melasma, chronic idiopathic urticaria, and alopecia areata. TCS are widely used for many other causes of skin inflammation.
Route of administration	Topical	Topical	Topical	Topical
Recommended dosage	Apply thin layer to affected areas twice daily.	Treatment: Applied morning and evening twice daily as a thin layer to affected areas of skin, including the face, neck, and	Treatment: Thin layer of pimecrolimus cream, 1% to sufficiently cover the affected skin area twice daily. May be used on all skin surfaces including the	Varies between drugs.

	Crisaborole	Tacrolimus	Pimecrolimus	Topical Corticosteroids
		eyelids. If no improvement occurs after 6 weeks of therapy or in case of disease exacerbation, therapy should be discontinued and patients should consult their physicians. Maintenance: Patients who have a high frequency of flares (≥ 5 times per year) and are responding to up to 6 weeks of acute treatment with tacrolimus ointment twice daily are suitable for maintenance treatment, applied once daily twice a week. There should be 2 to 3 days between applications (e.g., Monday and Thursday). Protopic should be applied as a thin layer to the areas of the skin normally affected by AD (including the face, neck, and eyelids). If flares recur, twice daily treatment should be reinitiated. In the absence of safety data for maintenance treatment beyond 12 months, a review of the patient`s condition should be conducted by the physician after 12 months of maintenance treatment and a decision taken whether to continue. In children, this review should include suspension of treatment to assess the need to continue the regimen and to evaluate the course of the disease.	head, neck, and intertriginous areas. Maintenance: Should be used for short or long intermittent periods of treatment. Therapy should be stopped upon clearance of the signs and symptoms of AD (e.g., pruritus, inflammation, and erythema). Treatment should be discontinued if resolution of disease occurs. If no improvement occurs after 3 weeks of treatment, or in case of disease exacerbation, therapy should be discontinued and patients should consult their physicians.	

	Crisaborole	Tacrolimus	Pimecrolimus	Topical Corticosteroids
Serious side effects / safety issues		Long-term safety of TCI has not been established. Although a causal relationship has not been established, rare cases of skin malignancy and lymphoma have been reported in patients treated with topical calcineurin.	Long-term safety of TCI has not been established. Although a causal relationship has not been established, rare cases of skin malignancy and lymphoma have been reported in patients treated with TCI.	If used under an occlusive dressing, particularly over extensive areas, or on the face, scalp, axilla(e), scrotum or when applied to the genitourinary tract, oral mucosa, or when administered rectally, sufficient absorption may take place to give rise to adrenal suppression and other systemic effects. Children may be at greater risk of developing systemic complications with the use of TCS.

AD = atopic dermatitis; PDE4 = phosphodiesterase type 4; TCI = topical calcineurin inhibitors; TCS = topical corticosteroids.

[a] Health Canada indication.

Source: e-CPS.[14]

Objectives and Methods

Objectives

To perform a systematic review of the beneficial and harmful effects of crisaborole ointment, 2% (Eucrisa) for the treatment of patients aged two years or older with mild-to-moderate AD.

Methods

All manufacturer-provided trials considered pivotal by Health Canada were included in the systematic review. Phase III studies were selected for inclusion based on the selection criteria presented in Table 3.

Table 3: Inclusion Criteria for the Systematic Review

Patient Population	Patients aged two years and older with mild-to-moderate AD **Subgroups:** • Severity (e.g., mild, moderate) • Failed or are intolerant to a topical corticosteroid treatment • Age (e.g., children, adults)
Intervention	Thin layer of crisaborole ointment 2% (Eucrisa) applied twice daily to affected skin areas.[a]
Comparators	**Topical corticosteroids** **Topical calcineurin inhibitors** • Pimecrolimus 1%[b] • Tacrolimus 0.03% and 0.1%[c] **Vehicle** • Placebo
Outcomes	**Efficacy outcomes:** • Severity of AD and AD lesions (e.g., ISGA score) • Symptom reduction (e.g., pruritus,[d] pain, sleep disturbance[d]) • HRQoL[d] (e.g., DLQI score, CDLQI score, DFI) • Exacerbations/flares[d] (e.g., need for additional health care utilization, hospitalization) • Mood (e.g., anxiety,[d] depression[d]) • Productivity (e.g., days of missed work/school) **Harms outcomes:** • AEs, SAEs, WDAEs, AEs of special interest (e.g., hypersensitivity)
Study Design	Published and unpublished phase III and IV RCTs

AD = atopic dermatitis; AE = adverse event; CDLQI = Children's Dermatology Life Quality Index; DFI = Dermatitis Family Impact; DLQI = Dermatology Life Quality Index; HRQoL = health-related quality of life; ISGA = Investigator's Static Global Assessment; RCT = randomized controlled trial; SAE = serious adverse events; WDAE = withdrawal due to adverse event.

[a] Crisaborole (Eucrisa) is for topical use only and not for ophthalmic, oral, or intravaginal use.

[b] Agent has Health Canada–approved indication for the treatment of patients two years of age and older with mild-to-moderate AD.

[c] Agent has Health Canada–approved indication for the treatment of patients with moderate-to-severe AD. Tacrolimus 0.03% and 0.1% are indicated for adults and only 0.03% is indicated for children aged 2 years to 15 years.

[d] Outcomes identified as important from patient input.

The literature search was performed by an information specialist using a peer-reviewed search strategy.

Published literature was identified by searching the following bibliographic databases: MEDLINE ALL (1946–) with in-process records and daily updates via Ovid; Embase (1974–) via Ovid; and PubMed. The search strategy consisted of both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concepts were Eucrisa (crisaborole).

No methodological filters were applied to limit retrieval by study type. Where possible, retrieval was limited to the human population. Retrieval was not limited by publication year or by language. Conference abstracts were excluded from the search results. See Appendix 2 for the detailed search strategies.

The initial search was completed on June 20, 2018. Regular alerts were established to update the search until the meeting of the CADTH Canadian Drug Expert Committee on October 17, 2018. Regular search updates were performed on databases that do not provide alert services.

Grey literature (literature that is not commercially published) was identified by searching relevant websites from the following sections of the Grey Matters checklist (https://www.cadth.ca/grey-matters): Health Technology Assessment Agencies; Health Economics; Clinical Practice Guidelines; Drug and Device Regulatory Approvals; Advisories and Warnings; Drug Class Reviews; Databases; and an Internet search. Google and other Internet search engines were used to search for additional Web-based materials. These searches were supplemented by reviewing the bibliographies of key papers and through contacts with appropriate experts. In addition, the manufacturer of the drug was contacted for information regarding unpublished studies.

Two CDR clinical reviewers independently selected studies for inclusion in the review based on titles and abstracts, according to the predetermined protocol. Full-text articles of all citations considered potentially relevant by at least one reviewer were acquired. Reviewers independently made the final selection of studies to be included in the review, and differences were resolved through discussion. Included studies are presented in Table 4; excluded studies (with reasons) are presented in Appendix 3.

Results

Findings From the Literature

A total of two studies were identified from the literature for inclusion in the systematic review (Figure 1). The included studies are summarized in Table 4. A list of excluded studies is presented in Appendix 3.

Figure 1: Flow Diagram for Inclusion and Exclusion of Studies

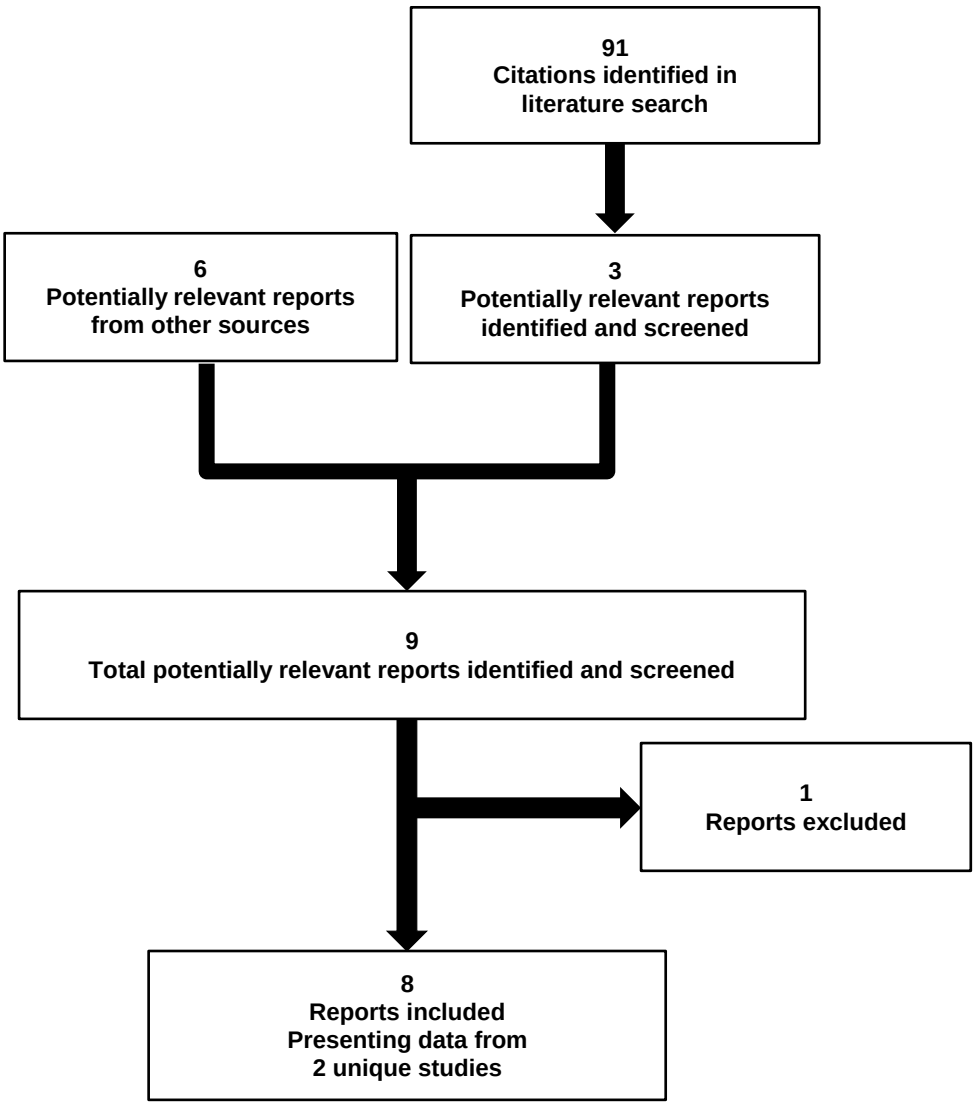


Table 4: Details of Included Studies

		AD-301	AD-302
Designs and Populations	**Study Design**	Phase III multi-centre, randomized, double-blind, vehicle-controlled study	Phase III multi-centre, randomized, double-blind, vehicle-controlled study
	Location	48 sites in the US	42 sites in the US
	Randomized (N)	763 (crisaborole [N = 507] and vehicle [N = 256])	764 (crisaborole [N = 514] and vehicle [N = 250])
	Inclusion Criteria	• Male or female aged 2 years and older at baseline/day 1 • Clinical diagnosis of AD according to Hanifin and Rajka • AD involvement ≥ 5% treatable % BSA, excluding the scalp • Had an ISGA score of mild (2) or moderate (3) at baseline/day 1	
	Exclusion Criteria	• Had any clinically significant medical disorder, condition, or disease or clinically significant physical examination finding at Screening that may interfere with study objectives • Had unstable AD or any consistent requirement for high potency TCS to manage AD signs and symptoms • Had a history of angioedema or anaphylaxis • Had a significant active systemic or localized infection, including known actively infected AD • Had a history of use of biologic therapy, including intravenous immunoglobulin, at any time prior to study • Had recent or anticipated concomitant use of systemic or topical therapies that might alter the course of AD, as specified in the protocol • Had undergone treatment for any type of cancer (except squamous cell carcinoma, basal cell carcinoma, or carcinoma in situ of the skin, curatively treated with cryosurgery or surgical excision only)	
Drugs	**Intervention**	Crisaborole 2% topical ointment applied b.i.d. on the skin only, excluding the scalp	
	Comparator(s)	Crisaborole vehicle (placebo) topical ointment applied b.i.d. on the skin only, excluding the scalp	
Duration	**Phase**		
	Run-in	Up to 35 days	
	Double-blind	28 days	
	Follow-up	7 days (day 36)	
Outcomes	**Primary End Point**	Proportion of patients achieving success in ISGA at day 29 (Success in ISGA was defined as ISGA of clear or almost clear with at least a 2-grade improvement from baseline)	
	Other End Points	**Secondary End Points**	
		• Proportion of subjects with an ISGA score of clear (0) or almost clear (1) at day 29 • Time to success in ISGA	
		Exploratory End Points	
		• Time to Improvement in Pruritus • Signs of AD (erythema, induration/papulation, exudation, excoriation, and lichenification) evaluated globally on a 4-point scale and not by body region • Change in CDLQI from baseline to day 29 for patients aged 2 to 15 years • Change in DLQI from baseline to day 29 for patients aged 16 years and older • Change in DFI from baseline to day 29 for parents/guardians of patients aged 2 to 17 years	
Notes	**Publications**	Paller 2016,[15] 2018[16]	

AD = atopic dermatitis; b.i.d. = twice daily; BSA = body surface area; CDLQI = Children's Dermatology Life Quality Index; CSR = clinical study report; DFI = Dermatitis Family Impact Questionnaire; DLQI = Dermatitis Life Quality Index; ISGA = Investigator's Static Global Assessment; TCS = topical corticosteroids.

Note: Six additional reports were included (FDA clinical[17] and statistical reviews;[18] manufacturer's submission;[19] and CSRs for AD-301[5] and AD-302[6]).

Source: CSRs for AD-301[5] and AD-302.[6]

Included Studies

Description of Studies

Two multi-centre, manufacturer-sponsored, double-blind RCTs met the inclusion criteria for this review. Studies AD-301 (N = 763; 48 sites) and 302 (N = 764; 42 sites) were identically designed and enrolled populations two years of age or older with mild-to-moderate AD (Investigator's Static Global Assessment [ISGA] scoring) comparing crisaborole in a 2:1 ratio to vehicle over a 28-day treatment course. Both studies were centred entirely in the US. The primary outcome was the proportion of patients with success by ISGA at day 29, while secondary outcomes included the proportion of patients with ISGA of clear or almost clear at day 29, and the time to success in ISGA.

After a screening period which lasted up to 35 days, patients were randomized using an interactive Web response system, stratified by study centre. The screening period allowed for washout of prior drug, but if patients did not require a drug washout they could start the study treatment immediately. After completing 28 days of study therapy, patients were to complete a post-treatment follow-up visit within seven days of day 28. Patients who completed their treatment and were eligible could be enrolled into an open-label extension (AD-303) if they wished to do so.

Populations

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria were similar for both the AD-301 and AD-302 trials. Patients aged two years and older were eligible for enrolment if they had an ISGA score of mild (2) or moderate (3) at baseline as well as a greater than 5% body surface area involvement. The diagnosis of AD was made in accordance with the Hanifin and Rajka criteria.[5,6] The criteria from Hanifin and Rajka consist of four basic features and 23 minor features that are indicative of a diagnosis of AD. Patients must have three or more basic features of: 1) pruritus, 2) morphological features (flexural lichenification in adults and facial and extensor eruptions in infants and children), 3) chronic or chronically relapsing dermatitis, and 4) personal or family history of atopy (asthma, allergic rhinitis, and AD), as well as three of 23 minor features. Minor features include a mix of morphological features (dry skin, chapped lips), laboratory values (elevated serum IgE), demographics (early age of onset), history (recurrent conjunctivitis, food intolerance), and symptoms (itch when sweating).[20]

Baseline Characteristics

The mean age across the two included studies was approximately 12 years, and the majority of patients were in the 2- to 11-year-old age range. There were more female (56% female) than male patients across both studies and the majority (approximately 60%) were white. All patients had an ISGA score of either mild or moderate, with the majority (approximately 61%) being moderate severity. There were no notable differences in baseline characteristics between groups.[5,6]

Table 5: Summary of Baseline Characteristics

Characteristics	AD-301 Crisaborole N = 503	AD-301 Vehicle N = 256	AD-302 Crisaborole N = 513	AD-302 Vehicle N = 250
Age (years)				
Mean (SD)	12.0 (11.64)	12.4 (10.70)	12.6 (12.65)	11.8 (12.56)
2 to 11 years, n (%)	317 (63.0)	151 (59.0)	310 (60.4)	164 (65.6)
2 to 6 years	162 (32.2)	78 (30.5)	173 (33.7)	93 (37.2)
7 to 11 years	155 (30.8)	73 (28.5)	137 (26.7)	71 (28.4)
12 to 17 years, n (%)	121 (24.1)	67 (26.2)	126 (24.6)	57 (22.8)
≥ 18 years	65 (12.9)	38 (14.8)	77 (15.0)	29 (11.6)
Sex, n (%)				
Male	219 (43.5)	113 (44.1)	231 (45.0)	112 (44.8)
Female	284 (56.5)	143 (55.9)	282 (55.0)	138 (55.2)
Race, n (%)				
American Indian or Alaska native	8 (1.6)	3 (1.2)	3 (0.6)	2 (0.8)
Asian	26 (5.2)	17 (6.6)	26 (5.1)	10 (4.0)
Black or African-American	138 (27.4)	61 (23.8)	147 (28.7)	78 (31.2)
Native Hawaiian or other Pacific Islander	0 (0.0)	4 (1.6)	7 (1.4)	4 (1.6)
White	308 (61.2)	162 (63.3)	309 (60.2)	144 (57.6)
Other	23 (4.6)	9 (3.5)	21 (4.1)	12 (4.8)
ISGA, n (%)				
0 (Clear)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
1 (Almost clear)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
2 (Mild)	196 (39.0)	93 (36.3)	197 (38.4)	100 (40.0)
3 (Moderate)	307 (61.0)	163 (63.7)	316 (61.6)	150 (60.0)
4 (Severe)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Mean (SD) treatable per cent BSA	18.8 (18.6)	18.6 (18.9)	17.9 (17.5)	17.7 (15.6)

BSA = body surface area; ISGA = Investigator's Static Global Assessment; SD = standard deviation.

Source: Clinical study reports for AD-301[5] and AD-302.[6]

Interventions

In both studies, crisaborole or vehicle (placebo) were applied twice daily to treatable areas (excluding the scalp), with enough ointment to cover each lesion. Treatment areas were designated and documented on day 1, and the patient or guardian were provided with this information about the designated treatment areas. If there were any new AD lesions then this documentation was updated, and treatment was to include these new lesions. The first application was performed by the investigators, and the patients and care providers were encouraged to participate in this initial application. Patients were told to wear loose fitting clothing, not wipe off study drug from the skin, avoid occlusion of the treated areas, and avoid swimming or bathing/washing the treated areas within four hours of application. Compliance was tracked using an electronic dosing diary. Dose modifications were not allowed during the study. Use of corticosteroids by any route was prohibited, as was the use of systemic immunosuppressants. Use of topical retinoids, benzoyl peroxide, or systemic antihistamines could not be escalated, decreased, or used on an as-needed basis (i.e., in a non-stable way).

Outcomes

The proportion of patients achieving success on the ISGA at day 29 was the primary outcome of AD-301 and AD-302, with success being defined as an ISGA of clear or almost clear and at least a 2-point improvement in scores from baseline. The ISGA employed in the included studies is a 5-point scale that provides a global clinical assessment of AD severity based on an ordinal scale, scored by an investigator or physician.[21] There is variability in the number of points included in the ISGA scale,[21] but the studies of interest for this review used a scale that ranges from 0 to 4.[5,6] A score of 0 corresponds to a grade of clear, 1 is almost clear, 2 is mild, 3 is moderate, and 4 is severe AD.[5,6] A 6-point scale exists, which reaches a maximum score of 5 (very severe); however, this was not used in the included studies. A decrease in score relates to an improvement in signs and symptoms. No minimal clinically important difference (MCID) was found for the ISGA.

Health-related quality of life (HRQoL) was assessed as an exploratory outcome using the Dermatology Life Quality Index (DLQI) (adults) and the Children's Dermatology Life Quality Index (CDLQI) (children), as well as the Dermatitis Family Impact questionnaire (DFI). The DLQI is a self-reported, 10-item questionnaire that refers to the preceding week and assesses six different aspects that may affect quality of life as a result of living with a dermatological condition.[22,23] The aspects included in the questionnaire are symptoms and feelings, daily activities, leisure, work or school, personal relationships, and side effects of treatment.[22,23] Each item is scored using a Likert scale that ranges from 0 to 3.[22,23] A score of 0, 1, 2, and 3 corresponds to the following descriptions of how much an aspect is affected by the disease, respectively: "not at all/not relevant," "a little," "a lot," and "very much."[23] The scores of each of the 10 items are summed for an overall DLQI score between 0 and 30 (or a percentage of 30).[22] The higher the score, the greater the impairment of quality of life. The MCID is a change in score of at least 3.3 from baseline. Further details regarding the validity of this instrument can be found in Appendix 5.

The CDLQI was developed using the same methods for the development of the DLQI, for children between the ages of 3 years and 16 years.[24] Like the DLQI, the CDLQI is a self-reported questionnaire that refers to the preceding week, used to assess the impact of skin disease on the quality of life but for children and may be completed with help from a parent or guardian. It also involves 10 questions that address the following topics: symptoms and feelings, leisure, school or holidays, personal relationships, sleep, and treatment. A score of 0,1,2, or 3 is assigned, respectively, to the following answers to each of the 10 questions: "not at all," "only a little," "quite a lot," or "very much."[24] No MCID was identified for the CDLQI. Further details regarding the validity of this instrument can be found in Appendix 5.

The DFI questionnaire was designed to assess the impact of disease on the quality of life of parents and families of children affected by AD.[25] It is a disease-specific, self-administered questionnaire that relies on a one-week recall, and consists of 10 items that were derived from ethnographical interviews and focus groups.[25] The 10 items of the questionnaire address the following topics: housework, food preparation, sleep, family leisure activity, shopping, expenditure, tiredness, emotional distress, relationships, and the impact on the carer's life due to helping with treatment.[25] Each question is scored on a four-point Likert scale ranging from 0 to 3, for an overall total ranging from 0 to 30 with a higher score corresponding to a greater negative impact on the family's quality of life due to AD.[25] No MCID was identified for the DFI. Further details regarding the validity of this instrument can be found in Appendix 5.

Studies AD-301 and AD-302 used a Severity of Pruritus Scale (SPS) to assess the extent or severity of itching in patients. It is scored on a 4-point numeric rating scale that ranges from 0 to 3. The ratings correspond to a grade and definition, where 0 is a grade of none or no itching; 1 is mild or occasional, slight itching/scratching; 2 is moderate or constant or intermittent itching/scratching which is not disturbing sleep; and 3 is severe or bothersome itching/scratching which is disturbing sleep. Once instructions had been provided, the scale was completed by the study participant or their parent or guardian, using an electronic diary, and based on a 24-hour recall.[5,6] The MCID was 0.20. Further details regarding the validity of this instrument can be found in Appendix 5.

Statistical Analysis

Sample Size Calculation

The sample size in each study was chosen to provide 90% power to achieve a statistically significant difference between groups at a statistical significance level of alpha equal to 0.05 for the primary end point, the proportion of patients who achieved success in ISGA, assuming a success rate of 20% with crisaborole and 10% with vehicle. This resulted in 500 patients as a target enrolment in the crisaborole group and 250 patients in the vehicle-treated group. The manufacturer also noted that to achieve that target sample of 750 patients, they would need to screen 1,000 patients, assuming a 33% screen failure rate.[5,6]

Statistical Analysis

The primary end point (proportion of patients achieving treatment success, defined an ISGA of clear or almost clear and at least a 2-point improvement in scores from baseline at day 29), and first secondary end point (patients with an ISGA of clear or almost clear at day 29) were analyzed using logistic regression with factors of treatment group and analysis centre, and was expressed as an odds ratio and 95% confidence interval (CI). Each site was expected to enrol a minimum of 12 patients; however, if this did not happen at a given site then data could be combined between principle investigators to achieve this minimum. The second secondary end point (time to success at ISGA [clear or almost clear with at least a two-grade improvement from baseline]) was analyzed using Kaplan–Meier analysis and the log-rank test. Patients who did not reach success at ISGA were censored at day 29. The first additional efficacy end point of time to improvement in pruritus was also analyzed using Kaplan–Meier. Descriptive statistics were used to analyze the second additional efficacy end point of signs of AD, as well as other end points such as DLQI.[5,6]

Missing Data

Missing values were derived using the Markov Chain Monte Carlo (MCMC) method. Briefly, the MCMC analysis followed these steps:

- The number of day 29 missing values that were to be estimated using MCMC were calculated (referred to as “nmiss”).
- A data set was created for each treatment group, with observed values and those requiring estimation by MCMC. Missing ISGA values were filled in using the MCMC method “5 x nmiss” times to generate “5 x nmiss” data sets. These data sets were then combined into one complete set for each group, by imputation.
- The percentage of patients achieving success for this primary outcome was then computed for each data set, and each complete data set was analyzed using logistic regression with factors for treatment group and analysis centre. The results from the

above analyses were then combined into a single inference using statistical analysis software.

The same methods were used to analyze the first secondary end point (ISGA of clear or almost clear) and second secondary end point (time to success at ISGA).[5,6]

Multiplicity

Multiple statistical testing was carried out in a hierarchical manner, where the first secondary end point was tested and needed to be declared statistically significant for second secondary end point to be declared statistically significant.[5,6] No other outcomes were included in the hierarchy.

Sensitivity Analyses

Two sensitivity analyses were performed, which were then combined into one analysis for each study. In the first sensitivity analysis, ISGA success was analyzed using a repeated-measures logistic regression with the dependent variable being ISGA success (dichotomized) and the independent factors: treatment, analysis centre, and visit. There was no imputation for missing data in this analysis. The second sensitivity analysis, again for dichotomous ISGA data, imputed missing data using a model-based multiple imputation method.

Subgroups

Subgroup analyses were performed for subgroups based on age (2 to 11 years; 12 to 17 years; and 18 years or older), ethnicity, and race. No interaction *P* values were reported.

Analysis Populations

In both included studies, the intention-to-treat (ITT) population included all patients randomized who were dispensed drug. The per-protocol population included all patients who completed day 29 visit without any major protocol violations, were adherent to medication (took 80% to 120% of study doses, and had not missed more than six consecutive doses).[5,6]

Patient Disposition

Patient disposition for the double-blind phase of the AD-301 and AD-302 trials are summarized in Table 6. Patients in the AD-301 and AD-302 trial were screened and enrolled in 48 and 42 sites in the US. In the AD-301 trial, a total of 925 patients were screened and 763 were randomized while a total of 923 patients in the AD-302 trial were screened and 764 were randomized. The manufacturer reported that the 162 screening failures in AD-301 and 159 in the AD-302 trial were the result of insufficient per cent body surface area or ISGA, voluntary withdrawal, or other reasons. Of the patients randomized in the AD-301 and AD-302 trials, 99.2% and 99.8% received at least one dose of crisaborole and 100% received at least one dose of placebo, respectively.

For the AD-301 and AD-302 trials, withdrawals were more common in the vehicle group (12.1% and 14.8%, respectively) as compared with the crisaborole group (5.9% and 6.0%, respectively). The manufacturer reported that the differences in withdrawals were primarily due to increases in withdrawals due to parent or guardian in the placebo group compared with the crisaborole group for both trials (7.0% versus 2.4% in AD-301 and 8.0% versus 2.7% in AD-302, respectively). The proportions of patients who withdrew as a result of AEs

were similar in the vehicle (1.4% and 0.8%, respectively) and crisaborole groups (1.0% and 1.6%, respectively).

Table 6: Patient Disposition

	AD-301		AD-302	
	Crisaborole	**Vehicle**	**Crisaborole**	**Vehicle**
Screened, N	925		923	
Randomized, N (%)	507	256	514	250
Received the intervention	503 (99.2)	256 (100)	513 (99.8)	250 (100)
Did not receive the intervention	4 (0.8)	0 (0)	1 (0.2)	0 (0)
Discontinued, N (%)	30 (5.9)	31 (12.1)	31 (6.0)	37 (14.8)
Adverse event	7 (1.4)	2 (0.8)	5 (1.0)	4 (1.6)
Withdrawal by subject	3 (0.6)	6 (2.3)	6 (1.2)	3 (1.2)
Withdrawal by parent/guardian	12 (2.4)	18 (7.0)	14 (2.7)	20 (8.0)
Lost to follow-up	5 (1.0)	4 (1.6)	4 (0.8)	4 (1.6)
Death	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Other	3 (0.6)	1 (0.4)	2 (0.4)	6 (2.4)
Intention-to-treat, N	503	256	513	250
Per-protocol, N	435	201	454	208
Safety, N	502	252	510	247

Source: Clinical study reports for AD-301[5] and AD-302.[6]

Exposure to Study Treatments

The mean plus or minus SD number of dosing days was 28.2 plus or minus 4.3 days with crisaborole and 26.8 plus or minus 6.7 days with vehicle in AD-301, and 27.8 plus or minus 4.6 days with crisaborole and 26.9 plus or minus 5.6 days with vehicle in AD-302.[5,6]

Use of concomitant medications related to AD was generally similar between groups within studies. Oral antihistamine use was most common (AD-301: 20% with crisaborole versus 22% with vehicle; AD-302: 28% with crisaborole versus 32% with vehicle), followed by topical anti-itch medications (anesthetics, antihistamines) in 16% of crisaborole and 20% of vehicle patients in AD-301, and 23% in each group in AD-302. Emollients (AD-301: 10% versus 14%; AD-302: 9% in each group) were also used. A small proportion of patients (AD-301: 3% in each group; AD-302: 3% with crisaborole and 6% with vehicle) also used TCS, despite this being prohibited per protocol.

Critical Appraisal

Internal Validity

Severity of AD and improvements in AD were measured with ISGA score by treating physicians or investigators. The validity of this score (range 0 to 4) for measuring mild-to-moderate severity, and judgment regarding what constitutes an improvement remain unknown, as did the MCID of the treatment effect as measured by the proportion of patients with change from baseline to day 29 in the score. Of note, there were remarkable vehicle (placebo) effects observed in both trials. There were 25% and 18%, respectively, in the two trials of patients who experienced success at the primary end point over the 29 days of the study period, which by definition was clear or almost clear with at least a 2-grade

improvement in ISGA score. Also, as per protocol, no other concomitant therapies including TCS were allowed and protocol violation was low (3%). This strong placebo effect could reflect the natural variation of disease symptoms over time, or it could reflect patients' response to the emollient effect of the vehicle itself, and emollients are used in managing mild AD. As a consequence, it could be highly unreliable to judge the difference of, for example, 1 to 2 or 3 to 2 in severity based on a score that is mostly a reflection of symptoms in an affected area. This is particularly problematic in assessing the treatment effect without considering the degree of change from baseline to day 29 in the grade of improvement. Therefore, a clear or almost clear with at least 2-grade improvement from baseline as the primary end point would be less prone to subjective bias comparing to a judgment on clear or almost clear alone, as one grade of difference would be difficult to discern (e.g., from 1 to 0).

The subjectivity is also an issue when patients suffered drug-related adverse effect, such as a higher incidence of application site pain (6.2% and 2.7% versus 1.2%, crisaborole versus vehicle, respectively, in both studies) than the vehicle group which might have led to the assessing physician becoming aware of the treatment assignment, and thus biased the estimate of ISGA score in favour of the study drug. Such situation may also apply to other drug-related adverse effects such as upper respiratory tract infection, pyrexia, nasopharyngitis, vomiting and nasal congestion, in which the incidence of the events was consistently higher in the crisaborole arm than the vehicle arm.

The manufacturer conducted calculations to determine the appropriate sample size for each of the included studies, and appeared to meet the minimum targets for enrolment in each group. The calculations were based upon assumptions of treatment success of 20% with crisaborole and 10% with vehicle, and it is not clear upon what basis these assumptions were made.

The manufacturer appeared to make adjustments for multiple statistical comparisons tests on the two secondary end points of ISGA success by applying a hierarchical testing procedure. Important outcomes such as HRQoL, assessed by the DLQI, and symptoms were not part of the statistical hierarchy and were not statistically assessed in either included study. Due to those random variability and possible subjective bias, the potential benefit of crisaborole on HRQoL improvement remains uncertain. AD can have a significant effect on quality of life, both in adults and children, thus this represents a significant gap in knowledge about crisaborole.

Subgroup analyses were planned and reported. No testing of heterogeneity on-treatment effect was reported. Additionally, there was no subgroup analysis reported based on baseline AD severity, although a post hoc analysis was provided to CADTH after a request of the manufacturer. Such subgroup analysis could be useful in revealing possible differential treatment effect by severity. In particular, it would help demonstrate the major driver (mild versus moderate AD) of treatment effect in the overall population.

The manufacturer has requested that crisaborole be reimbursed for patients who have failed or are intolerant to TCS, yet no subgroup analyses were provided assessing whether responses differed based on these subpopulations.

The included studies were both double-blinded; this was accomplished by the use of vehicle control. There were numerically more withdrawals in the vehicle-treated group than in the crisaborole group, and the difference seems to have been largely accounted for by withdrawals by parent or guardian. It is not clear whether this difference is an indication of a

problem with the blinding. Unlike a typical matched control for an orally administered drug, a topically applied vehicle needs to be matched in all respects, including texture and odour, presenting additional challenges. It is not clear whether the manufacturer matched vehicle in all these respects. Additionally, there was a numerical increase in topical AEs with crisaborole such as application site pain, and this might have also made it difficult to maintain blinding in these patients.

The manufacturer used a MCMC method to impute missing data, and this method assumes data are missing at random. The fact that more patients withdrew in the vehicle group than in the crisaborole group in both studies, and that most of this difference was accounted for by lack of efficacy might indicate that an assumption of data missing at random might not be appropriate. Of note, the FDA statistical reviewer conducted a sensitivity analysis assuming all missing values constituted treatment failures and this did not change the overall conclusions on the relative effect of crisaborole versus vehicle on the primary end point.[18] It is likely that missing data on the primary end point due to withdrawal was small and therefore did not have a substantial impact on the findings.

The use of concomitant TCS was prohibited in the included studies; however, there were a small percentage of patients in each of the included studies that were reported as using TCS during the study. These represent a protocol violation and potentially an important one as TCS are considered the standard therapy for treating AD. The proportion of patients using TCS was small, however; 3% in each group in AD-301 and 3% with crisaborole and 6% with vehicle in AD-302, and this may have mitigated the impact on the overall analysis.

External Validity

The populations in both studies were generalizable to the Canadian population that might use crisaborole with respect to age and sex, according to the clinical experts consulted for this review. However, both studies were conducted entirely in the US, and there was a relatively high proportion of African-Americans enrolled in each study, compared with the proportion one would expect to see in Canada. The clinical experts also noted that instruments relying on visual assessment of AD may be less reliable in patients with darker skin as changes in colour and other morphology may be more challenging to detect. The clinical experts also noted the relatively high proportion of responders in the vehicle-treated group, particularly in AD-301 (treatment success in 25% of patients), and speculated as to whether this high vehicle response might have been at least in part due to difficulties in assessing patients with more pigmented skin tones.

The primary outcome in the included studies was based on the ISGA, which relies on investigator assessment, in this case looking for responses of clear or almost clear on AD lesions. Thus this is a subjective measure; however, it is widely accepted as a standard instrument for assessment of AD in clinical trials according to the clinical experts consulted on this review. Yet, this ISGA score was not used in routine clinical practice as a measure to judge the improvement in treatment effect. This would make it uncertain whether the observed treatment effect as measured by this score could be readily interpretable to and meaningful in a real-world setting. The time frame for detecting an improvement appears short at 29 days; however, the clinical experts also believed this to be adequate for detecting a response to therapy.

There was no active comparator in either of the included studies. The two most relevant comparators would be the TCI and the TCS. The manufacturer-requested reimbursement criteria includes patients who are intolerant to or have failed TCS, thus, the efficacy and

harms of crisaborole relative to TCS are unknown. TCI are typically considered to be an alternative to TCS in these unresponsive or intolerant patients, or in patients who wish to avoid using TCS due to their side effects, thus the lack of comparative data versus the TCI is a significant gap in knowledge.

Crisaborole employs a unique mechanism of action for topically applied drugs and is first in its class; therefore it is important to have long-term safety data regarding this drug. The included studies double-blind phase ended after 29 days, therefore this is not of sufficient duration to determine if there are any long-term safety issues associated with the use of this drug. AD is a chronic condition and patients would be expected to use a drug like crisaborole for long periods of time. There is data available from a longer-term extension; however, there is no longer a control group in this phase of the study.

Efficacy

Only those efficacy outcomes identified in the review protocol are reported below.

Severity of Atopic Dermatitis and Atopic Dermatitis Lesions

Proportion of Patients Achieving Success in Investigator's Static Global Assessment at Day 29

The primary outcome in both trials was the proportion of patients achieving success in ISGA at day 29. Success was defined as ISGA of clear or almost clear with at least a 2-grade improvement from baseline/day 1. For both trials, patients treated with crisaborole were statistically significantly more likely to have success as compared with those treated with vehicle (AD-301 = 32.8% versus 25.4%; *P* value: 0.038 and AD-302 = 31.4% versus 18.0%; *P* value < 0.001) (Table 7). The proportion of vehicle-treated patients who had a success was higher in the AD-301 trial as compared with those in the AD-302 trial (25.4% versus 18.0%). In addition, the manufacturer conducted an analysis using the per-protocol population, which consisted of 83% of the ITT population in Study 201 and 87% of the ITT in Study 202. [redacted]

Subgroup analyses were presented based on age for the primary outcome, however, no interaction *P* values were reported and no formal comparisons were conducted between treatment groups. Nevertheless there was no clear indication that patient age impacts response to crisaborole. Additionally, in response to a request from CDR, a post hoc subgroup analysis with data pooled from both included studies was presented with responses broken down by baseline AD severity (ISGA, mild versus moderate). In this case the treatment effect was not statistically significant in patients with mild AD (success in 24.9% of crisaborole patients versus 21.1% of vehicle patients; mean difference between groups of 3.6 [95% CI, –3.9 to 11.2; *P* = 0.35]), while it was statistically significant in patients with moderate disease (36.7% versus 22.3% respectively; mean difference between groups of 14.4% [95% CI, 8.0 to 20.8; *P* < 0.0001]).[26] No interaction *P* values were reported.

Proportion of patients achieving ISGA of clear or almost clear (treatment success) was a secondary outcome of both studies. The proportion of patients with success was higher with crisaborole than vehicle in studies AD-301 (51.7% versus 40.6%, *P* = 0.005) and AD-302 (48.5% versus 29.7%, *P* < 0.001). Additionally, median time to treatment success was

reported, and although these values could not be calculated for the vehicle group (had not reached a 50% response) the differences between groups were also reported as statistically significant (Table 7).

Health-Related Quality of Life

Dermatology Life Quality Index

HRQoL was assessed using the DLQI in both studies, both in adults (DLQI) and children (CDLQI). In AD-301, the mean (SD) decrease (improvement) from baseline to day 29 was –5.5 (5.5) for crisaborole and –3.6 (4.6) for vehicle (Table 7). In AD-302 the mean (SD) decrease from baseline to day 29 was –5.0 (5.5) for crisaborole and –3.4 (5.8) with vehicle. With respect to the children's DLQI, the mean (SD) reduction from baseline to day 29 in AD-301 was –5.2 (5.6) with crisaborole and –3.1 (5.9) with vehicle, and in AD-302 was –4.0 (4.9) with crisaborole and –2.9 (5.0) with vehicle.

Dermatology Family Impact Questionnaire

The mean (SD) reduction (improvement) from baseline to day 29 in the Dermatology Family Impact Questionnaire (DFI) in AD-301 was –3.9 (5.7) with crisaborole and –2.7 (5.6) with vehicle, and in AD-302 it was –3.6 (5.2) with crisaborole and –2.8 (4.8) with vehicle (Table 7).

Symptoms

Time to Improvement in Pruritus

The median time to improvement in pruritus was an exploratory end point in both included studies. The median time to improvement in pruritus was 1.32 days with crisaborole and 1.87 days with vehicle ($P < 0.001$) in AD-301, and 1.41 days with crisaborole and 1.54 days with vehicle ($P = 0.425$) in AD-302 (Table 7).

Table 7: Key Efficacy Outcomes

Efficacy Outcomes	AD-301		AD-302	
Primary End Points	**Crisaborole N = 503**	**Vehicle N = 256**	**Crisaborole N = 513**	**Vehicle N = 250**
Proportion of Patients Achieving Success in ISGA at Day 29[a]				
Baseline ISGA				
2 (Mild)	196 (39.0%)	93 (36.3%)	197 (38.4%)	100 (40.0%)
3 (Moderate)	307 (61.0%)	163 (63.7%)	316 (61.6%)	150 (60.0%)
Success[a]	165 (32.8%)	65 (25.4%)	161 (31.4%)	45 (18.0%)
Failure	338 (67.2%)	191 (74.6%)	352 (68.6%)	205 (82.0%)
P value[b]	$P = 0.038$		$P < 0.001$	
Secondary Outcomes				
ISGA of Clear or Almost Clear at Day 29				
Success[a]	260 (51.7%)	104 (40.6%)	249 (48.5%)	74 (29.7%)
Failure	243 (48.3%)	152 (59.4%)	264 (51.5%)	176 (70.3%)
P value	$P = 0.005$		$P < 0.001$	
Time to Success in ISGA[a] (Days)				
N	503	256	513	250
Median	NA	NA	NA	NA

Efficacy Outcomes	AD-301		AD-302	
P value	*P* < 0.001		*P* < 0.001	
Exploratory End Points				
Time to Improvement in Pruritus (Days)				
N	428	210	439	211
Median	1.32	1.87	1.41	1.54
P value	*P* < 0.001		*P* = 0.425	
Dermatology Life Quality Index (DLQI)				
Baseline N	95	52	97	40
Baseline mean (SD)	9.6 (6.37)	9.5 (6.52)	9.7 (6.24)	9.1 (6.67)
Mean (SD) change from baseline to day 29	–5.5 (5.45) N = 87	–3.6 (4.60) N = 44	–5.0 (5.49) N = 93	–3.4 (4.75) N = 38
Children's Dermatology Life Quality Index (CDLQI)				
Baseline N	393	199	404	204
Baseline mean (SD)	9.7 (6.19)	9.1 (6.54)	9.0 (5.77)	8.9 (5.48)
Mean (SD) change from baseline to day 29	–5.2 (5.63) N = 374	–3.1 (5.90) N = 175	–4.0 (4.92) N = 376	–2.9 (5.01) N = 180
Dermatitis Family Impact Questionnaire (DFI)				
Baseline N	431	214	431	217
Baseline mean (SD)	8.5 (6.63)	7.5 (6.66)	7.7 (6.57)	8.0 (5.65)
Mean (SD) change from baseline to day 29	–3.9 (5.68) N = 407	–2.7 (5.61) N = 187	–3.6 (5.18) N = 404	–2.8 (4.75) N = 190
Treatable % BSA				
Baseline mean (SD)	18.8 (18.55) N = 503	18.6 (18.87) N = 256	17.9 (17.49) N = 513	17.7 (15.61) N = 250
Mean (SD) change from baseline to day 29	–8.2 (12.83) N = 477	–5.8 (12.79) N = 228	–6.7 (12.22) N = 486	–3.1 (11.07) N = 224

BSA = body surface area; CLDQI = Children's Dermatology Life Quality Index; DFI = Dermatitis Family Impact Questionnaire; DLQI = Dermatology Life Quality Index; ISGA = Investigator's Static Global Assessment; NA = not available; SD = standard deviation.

[a] Success in ISGA for the primary outcome was defined as ISGA of clear or almost clear with at least a 2-grade improvement from baseline/day 1. This differed from the secondary end point, which defined success as simply being clear or almost clear.

[b] The *P* value from a logistic regression (with Firth option) test with factors of treatment group and analysis centre. The adjusted estimates for AD-301 and AD-302 from logistic regression were 29.1% and 22.0% and 26.5% and 14.2% for the crisaborole and vehicle groups, respectively. Values were adjusted for multiple imputation.

Improvement in pruritus defined as achieving none (0) or mild (1) with at least a 1-grade improvement from baseline. Medians computed using Kaplan–Meier methods.

Source: Clinical study reports for AD-301[5] and AD-302.[6]

Harms

Only those harms identified in the review protocol are reported below (see 2.2.1, Protocol). See Table 8 for detailed harms data.

Adverse Events

AEs were reported in 29% of patients in the crisaborole group and 20% of patients in the vehicle group in AD-301, and in 29% of crisaborole and 30% of vehicle-treated patients in AD-302 (Table 8). Application site pain was the most common AEs in AD-301, occurring in 6.2% of crisaborole-treated versus 1.2% of vehicle-treated patients after 29 days. In AD-302, application site pain occurred in 2.7% of crisaborole-treated and 1.2% of vehicle-treated patients after 29 days.[5,6]

Serious Adverse Events

There were few serious adverse events (SAEs) through the 29-day treatment period of either study. In AD-301, 1.0% of crisaborole-treated and 0.4% of vehicle-treated patients had an SAE, while in AD-302 0.6% of crisaborole-treated patients experienced an SAE, and none in the vehicle group (Table 8).[5,6]

Withdrawals Due to Adverse Events

In AD-301, withdrawal due to adverse event (WDAE) occurred in 1.4% of crisaborole-treated patients and in 0.8% of vehicle-treated patients, and in AD-302 WDAE occurred in 1.0% of crisaborole patients and 1.6% of vehicle-treated patients. The only WDAE that occurred in more than one patient in either study was application site pain and application site urticaria, each occurring in two patients in the crisaborole group in AD-301 (Table 8).[5,6]

Mortality

There were no deaths in either study.[5,6]

Notable Harms

One patient had a hypersensitivity reaction, reported as a treatment-emergent AE, in the crisaborole group in AD-302. Otherwise there were no other hypersensitivity reactions reported.[5,6]

Table 8: Harms

	AD-301		AD-302	
Adverse Events	**Crisaborole N = 503**	**Vehicle N = 256**	**Crisaborole N = 513**	**Vehicle N = 250**
Subjects with > 0 AEs, n (%)	147 (29.3)	50 (19.8)	150 (29.4)	79 (32.0)
Most common AEs (≥ 1% of patients)				
Application site pain	31 (6.2)	3 (1.2)	14 (2.7)	3 (1.2)
Upper respiratory tract infection	14 (2.8)	10 (4.0)	16 (3.1)	5 (2.0)
Pyrexia	12 (2.4)	3 (1.2)	7 (1.4)	4 (1.6)
Nasopharyngitis	9 (1.8)	0 (0.0)	9 (1.8)	6 (2.4)
Vomiting	8 (1.6)	3 (1.2)	7 (1.4)	2 (0.8)
Nasal congestion	7 (1.4)	0 (0.0)	–	–
Application site pruritus	4 (0.8)	3 (1.2)	1 (0.2)	3 (1.2)
Diarrhea	–	–	6 (1.2)	2 (0.8)
Application site urticaria	–	–	0 (0.0)	3 (1.2)
Staphylococcal skin infection	–	–	1 (0.2)	4 (1.6)
Headache	–	–	6 (1.2)	1 (0.4)
Cough	–	–	7 (1.4)	7 (2.8)
Oropharyngeal pain	–	–	7 (1.4)	2 (0.8)
Dermatitis atopic	–	–	4 (0.8)	6 (2.4)
Eczema	–	–	3 (0.6)	3 (1.2)
Pruritus	–	–	4 (0.8)	3 (1.2)
Serious Adverse Events				
Subjects with > 0 SAEs, n (%)	5 (1.0)	1 (0.4)	3 (0.6)	0

	AD-301		AD-302	
Appendicitis	1 (0.2)	0	–	–
Application site infection	–	–	1 (0.2)	0
Asthma	1 (0.2)	0	–	–
Cellulitis	0	1 (0.4)	–	–
Kawasaki disease	1 (0.2)	0	–	–
Laceration	–	–	1 (0.2)	0
Pneumonia	1 (0.2)	0	–	–
Suicide attempt	1 (0.2)	0	–	–
Suicidal ideation	–	–	1 (0.2)	0
Withdrawals Due to AEs				
WDAEs, n (%)	7 (1.4)	2 (0.8)	5 (1.0)	4 (1.6)
Most common reasons				
Application site pain	2 (0.4)	0	1 (0.2)	0
Application site urticaria	2 (0.4)	0	–	–
Application site vesicles	0	1 (0.4)	–	–
Impetigo	1 (0.2)	0	–	–
Pharyngitis	1 (0.2)	0	–	–
Dermatitis atopic	0	1 (0.4)	0	1 (0.4)
Eczema	–	–	1 (0.2)	0
Henoch-Schönlein purpura	–	–	0	1 (0.4)
Photosensitivity reaction	–	–	0	1 (0.4)
Swelling face	–	–	0	1 (0.4)
Kawasaki disease	1 (0.2)	0	–	–
Deaths				
Number of deaths, n (%)	0	0	0	0
Notable Harms				
Hypersensitivity reactions, n (%)	0	0	1 (0.2)	0

AE = adverse events; SAE = serious adverse events; WDAE = withdrawals due to adverse events.

Source: Clinical study reports for AD-301[5] and AD-302.[6]

Discussion

Summary of Available Evidence

Two multi-centre, manufacturer-sponsored, double-blind RCTs met the inclusion criteria for this review. Studies AD-301 (N = 763) and 302 (N = 764) were identically designed and enrolled populations two years of age and older with mild-to-moderate AD (based on ISGA scoring) comparing crisaborole in a 2:1 ratio to vehicle over a 28-day treatment course. The primary outcome was the proportion of patients with success by ISGA at day 29, while secondary outcomes included the proportion of patients with ISGA of clear or almost clear at day 29, and the time to success in ISGA.

HRQoL was not assessed formally in either included study; therefore, no conclusions can be drawn about the impact of crisaborole on quality of life, an important consideration for patients with AD. There were numerically more withdrawals in the vehicle group than with crisaborole, and this difference appears to have been largely accounted for by lack of efficacy. The included trials lacked an active comparator, therefore the relative efficacy of crisaborole to TCS or TCI is unknown. The included studies had a relatively short duration of follow-up and thus long-term efficacy and safety of crisaborole is unknown.

Interpretation of Results

Efficacy

The manufacturer-requested reimbursement criteria for crisaborole are for patients who have failed or who are intolerant to TCS treatment. The treatment effect of crisaborole compared with vehicle from the included studies was modest at best, thus its role presumably will be to provide a safer alternative to TCS or TCI. Although statistically significant, the difference in treatment success (primary outcome) between crisaborole and vehicle was between 7% and 13%, and the proportion of vehicle-treated patients with success was relatively high (18% to 25%) across the two studies. TCS have traditionally been first-line therapies for AD; however, their use has always been somewhat limited by their adverse effects, as their efficacy is well established. Corticosteroids in general have a reputation for safety issues and these are primarily derived from their use as systemic drugs, rather than topically. Within the dermatology community it appears to be generally accepted that TCS are safe and effective when used wisely, with careful attention paid to location, potency, and duration of use.[27] This disconnect between the actual and perceived harms associated with TCS is so well established that it has been given the name, steroid phobia, and has been well studied in the literature. A recent systematic review by Li et al. found 16 studies that assessed the issue of steroid phobia in patients or caregivers, in AD.[28] The authors found a wide range in prevalence of steroid phobia between studies (21% to 84% of patients). This is likely attributable at least in part to a wide range of definitions of steroid phobia. There were two studies that assessed the relation between steroid phobia and patient adherence, and those found that patients with steroid phobia had a significantly higher risk of nonadherence compared with those without fear of using TCS (49% versus 14% in one study, respectively, and 29% versus 10% in the other). Nonadherence may certainly contribute to failure of patients to respond to TCS, and. of course, lack of response may lead patients to try an alternative therapy, the most likely alternative being TCI. The TCI can thus be considered an important comparator for crisaborole. TCI have their own reputation for adverse effects, as they carry a safety

warning related to the potential for increased risk of lymphoma and skin cancer. Although the safety warning comes with a number of caveats, namely being that it is far from well established, it would likely be a significant consideration for users of TCI, including caregivers (parents) who are considering whether to use TCI on their children. Thus one can see where crisaborole may potentially become a popular option for AD, as it currently lacks the reputation for causing harms that are seen with TCS and TCI. That said, patient education and reassurance would likely enhance patient willingness to try TCS or TCI.

There were no studies that compared crisaborole to an active comparator, most notably a TCS or a TCI. There was also limited information available from indirect comparisons that might further inform the relative efficacy or harms of crisaborole to these comparators. Two network meta-analyses (NMA) were found in the literature; however, both were limited by a lack of available trials (for further details, see Appendix 7). There were no comparisons to TCS and a limited number of trials in which to compare crisaborole with the TCI (pimecrolimus and tacrolimus). There were no statistically significant differences found between crisaborole and TCI for patients achieving an ISGA of clear or almost clear, and there was no NMA performed for safety outcomes, and no subgroup analyses were available due to the lack of data. Therefore with respect to either direct or indirect comparisons versus active comparators, the relative efficacy and harms of crisaborole compared with TCS and TCI is unknown.

Crisaborole is approved for use in patients with mild-to-moderate AD; however, there were no pre-planned subgroup analysis results available looking at responses broken down by baseline severity of AD. In response to a request from CDR, the manufacturer provided a post hoc subgroup analysis of primary outcome responses by baseline ISGA (mild versus moderate).[26] The results suggest there is a relatively modest treatment effect that failed to reach statistical significance in patients with mild disease at baseline compared with those with moderate disease, where there was a clear statistically significant difference between groups. This is not necessarily a surprising finding, as patients with mild AD have limited room for improvement; however, clinically it does suggest that patients with moderate disease are more likely to benefit from treatment. This observation is supported by patient input to CDR, which suggests that patients with moderate-to-severe disease are most impacted by their condition. Therefore, the differential response is not unexpected and the results, particularly the lack of statistical significance in the mild group was most likely due to a small difference on the change from baseline and most likely cannot be attributable to a power issue. In contrast, the moderate group with similar (though slightly larger) sample size showed statistical significance, most likely due to a large difference in the change between treatment groups.

The lack of assessment of HRQoL data from the included studies is a limitation of this review. In their input to CDR, patients identified a number of quality of life issues associated with AD including persistent itch. Without a formal assessment of quality of life there is no way to ascertain whether crisaborole improves this important outcome. The MCID for the DLQI is 3.3, and both the crisaborole and vehicle groups achieved an improvement from baseline that exceeded the MCID. It is not clear whether the MCID also applies to between-group differences, thus the impact of crisaborole on the DLQI is unknown. There is no MCID for the CDLQI, and due to the relatively large proportion of patients under the age of 18 in both studies, much more data were available for this outcome than the DLQI. The lack of study of this outcome, as well as the DFI, means that no conclusions can be drawn about the impact of crisaborole on HRQoL in children.

Harms

The most common AEs suggest that there may be some tolerability issues associated with the use of crisaborole, most notably application site pain. There was no indication that WDAE would be higher with crisaborole than vehicle. If there are application site issues like pain associated with crisaborole it is not clear what the reason would be, as this is the first phosphodiesterase type 4 (PDE4) inhibitor to be approved for topical use. Given that this was a vehicle-controlled study, it is likely that the issue is attributable to the drug and not any excipients in the ointment. No longer term safety issues were identified from the extension study, AD-303; however, this was an uncontrolled extension and thus limits any conclusions that can be drawn regarding safety.[29]

Potential Place in Therapy[2]

Based on current therapies and standard of care for AD, an unmet need would be effective, affordable, and safe therapies for patients suffering with the disease. One challenge would be steroid phobia. While this would best be addressed through proper patient education about appropriate use of topical steroids, crisaborole would add to the armamentarium of treatment options for these patients. Another challenge would be topical treatment options for patients with unresponsive and/or severe AD. In this respect, it is unknown whether or not crisaborole would meet this need. With respect to affordability, the projected cost of approximately $2.00 per gram of medication makes it comparable with TCI and does not offer an advantage as compared with TCS.

There are some fears about the black box labelling of TCI. While this labelling may generate more fear in the patient population about using TCI, the real-world safety is likely not impacted by switching to crisaborole. Nonetheless, there may be a perceived safety by patients in using crisaborole.

Patients with mild-to-moderate AD who are unwilling to use topical steroids (e.g., for reasons such as steroid phobia) would be good candidates to receive topical crisaborole. No special diagnostic tests would need to be run.

Conclusions

Two identically designed multi-centred, double-blind RCTs (AD-301 and AD-302), both entirely based in the US, met the inclusion criteria for this systematic review. Both studies randomized patients with AD scored as mild-to-moderate using ISGA, in a 2:1 ratio to either crisaborole or vehicle over a treatment course of 28 days. A larger percentage of crisaborole-treated patients versus vehicle achieved the primary outcome, treatment success according to the ISGA at day 29, and this difference was statistically significant in both studies. No conclusions can be drawn regarding HRQoL, and this is an important limitation given the impact of AD on this outcome. There was some indication of a numerically higher risk of application site pain with crisaborole, although there was no increased risk of WDAE. The lack of an active comparator in the two trials was a limitation of this review. The data available from the NMA suggests no statistically significant differences for crisaborole versus TCI for the percentage of patients achieving an ISGA of clear or almost clear; however, there were several limitations with the analysis, and no data were available to compare crisaborole to TCS.

[2] This information is based on information provided in draft form by the clinical expert consulted by CDR reviewers for the purpose of this review.

Appendix 1: Patient Input Summary

This section was prepared by CADTH staff based on the input provided by patient groups.

1. Brief Description of Patient Group(s) Supplying Input

Two patient groups responded to the call for patient input for this CDR review.

One of the groups was the Canadian Skin Patient Alliance (CSPA), which is a non-profit organization that serves patients with dermatological conditions, diseases, and traumas in Canada. It focuses on education and advocacy for these patients as well as more than 20 additional affiliated disease-specific organizations in Canada.

The CSPA declared the following companies have provided their group with financial payment in the last two years and may have an interest in the drug under review: Pfizer, Novartis, Janssen, Galderma, and AbbVie. The patient group also declared that the National Eczema Association in the US assisted with the distribution of a survey that was highlighted in their Patient Input Summary. Lastly, there were no conflicts of interests to report regarding the preparation of this submission.

The Eczema Society of Canada (ESC) responded to the call for patient input as well. The ESC is a registered Canadian charity dedicated to improving the lives of Canadians living with atopic dermatitis (AD) or eczema. Expert health care professionals assist the ESC with the delivery of up-to-date information regarding the disease and treatment information to Canadians living with AD, which includes patients and caregivers, as well as health care providers.

The ESC declared the following companies have provided their group with financial payment in the last two years and may have an interest in the drug under review: Actelion Pharmaceuticals Ltd., Beiersdorf Canada Inc., Bioderma Canada, Blistex Inc., Familiprix Canada, Galderma Canada Inc., GlaxoSmithKline Canada, Johnson & Johnson Inc., Leo Pharma Inc. Canada, L'Oréal Canada Inc., Paladin Labs Inc., Pierre Fabre Dermo-Cosmétique Canada Inc., Pfizer Global Inc., Pfizer Canada Inc., Sanofi Consumer Health, Sanofi Genzyme Canada, Unilever Canada, Bausch Health Canada (formerly Valeant Canada), and WellSpring Pharmaceuticals. They also reported funding from private citizen donations as well as the following organizations/grant programs: CanadaHelps, FedEx Cares Employee Community Fund, and IBM Canada. The patient groups declared no conflict of interests in the preparation of this submission.

2. Condition-Related Information

Survey data were used to inform the patient input response prepared by both of the patient groups.

The CSPA conducted two surveys: the first was completed in Canada and focused on gaining an understanding of how AD affects individuals and their caregivers; the second was completed in the US to gain insight into the patient experience with crisaborole, as none of the clinical trials were conducted in Canada. A total of 194 and 28 respondents completed the first and second surveys, respectively. Of the 194 participants in the first survey, 132 were patients living with the disease and 62 were caregivers for children living with the disease. The majority (78%) of respondents were female, with ages ranging from 18 to 92 years old. The second survey conducted in the US included 28 respondents, 25 of whom had experience with crisaborole. Additionally, two of the respondents who had experience with crisaborole identified as caregivers for children with AD.

The ESC surveyed Canadians living with AD via online surveys, with the intent of gaining a better understanding of the burden of disease for this population, and the barriers that exist and hinder better care. A total of 1,035 respondents completed the survey, which included respondents from each of the provinces across Canada. Adults living with AD and their caregivers made up 36% of the respondents; the remaining 64% were children living with the disease and their caregivers. None of the respondents to the ESC survey had experience with crisaborole.

The patient groups defined AD, more commonly known as eczema, as a chronic, inflammatory skin condition. According to the CSPA, up to 20% of adults and 9% of children are affected by AD in Canada. AD can be classified as mild, which is characterized by areas of dry skin and infrequent itching, as well as small areas of redness in some cases. It can also be classified as moderate/severe, which involves areas of dry skin, more frequent itching, and redness. Some patients also experience localized skin thickening and lesions that can ooze and bleed during flare-ups.

Based on the CSPA's survey, patients with mild AD reported experiencing a minor overall impact, with some not even using medicated treatment. However, moderate-to-severe AD has a greater impact on the lives of patients and their caregivers, with reports of regular interrupted sleep, negative effects on work and school life, as well as personal life. One patient stated, "I am very sensitive to artificial fragrances (perfumes, candles, air fresheners) and can only use products with no scents. I had to leave high school because of my sensitivity and attend school online." The patient group also suggested that AD may be related to other comorbidities, as many of the respondents also suffer from seasonal and/or food allergies. Further, AD may also increase the patient's susceptibility to infection and complication, with 38% of respondents reporting an episode involving oozing lesions as a result of Staphylococcus aureus infection in the past. A patient noted that they "had continuous flare ups over the course of 7 years. From head to toe. Red, weeping, sores, burning, lots of pain. It has contributed to my depression heavily and made working and living very hard." This information was echoed by the ESC patient group as well.

Living with AD also has an effect on the mental well-being of individuals. For example, the survey conducted by the ESC noted the following issues reported by adult patients with moderate AD: anxiety related to AD (61%), avoidance of social activities (40%), avoidance of exercise and physical activity (33%), depression related to AD (32%), avoidance of intimacy (26%), missed work and/or important life events (26%), and change of career or give up certain activities due to the disease (23%).

"People often underestimate this disease but it is all consuming and so hard to manage. Sometimes I can't even think straight because I'm so itchy. I have to work, be a mom, and function all while feeling so irritable due to lack of sleep and itchy skin."

The ESC reported that based on the 384 survey responses pertaining to children who live with mild or moderate AD, 25% experienced sleep loss during eight nights of the month due to the disease, 24% miss at least 10 days of school a year due to the disease, 21% have difficulty participating in sports or physical activities, 7% are bullied and/or picked on by peers, and 4% experience depression due to their AD. Dealing with AD also has an impact on the lives of parents and caregivers, with 44% reporting sleep loss due to their child's AD and 27% experience anxiety specifically related to their child's disease. This is also reflected in the submission from CSPA, which highlights 68% of caregivers reported that AD has a negative impact on their lives, with comments from caregivers regarding the care for children with AD being overwhelming at times.

3. Current Therapy-Related Information

According to the response from both patient groups, finding adequate treatment for AD is a challenge for many patients. Based on the survey data from both patient groups, the majority of patients had tried multiple treatments in an effort to control their condition, with between 23% and 31% of respondents reported having tried 10 or more treatments to manage their AD. As one patient describes, “I haven’t found a cure or perfect routine. One thing works for months and then suddenly it doesn’t anymore and you have to start over.” The response from both groups indicated that TCS, such as hydrocortisone, betamethasone, and clobetasol, were the most commonly used available treatments for both mild and moderate-to-severe AD; however, the CSPA group noted that “it is not a cure, and flares do return.” The second most common method of treatment, as reported by both patient groups, was the use of non-medicated methods such as skin care maintenance (bathing and moisturizing techniques). Other methods of treatment described in the CSPA response include: natural or herbal remedies, acupuncture, ultraviolet light A or B phototherapy, oral corticosteroids, methotrexate, azathioprine, antihistamines, TCI, and cyclosporine A.

The CSPA group also highlighted skin thinning from medication, spider veins, and blistering as common adverse events (AEs) for those living with mild AD. Similar AEs were reported for those living with moderate-to-severe AD, in addition to headaches. A total of 48% of respondents to the ESC patient input said treatment was uncomfortable, 47% said it was difficult to dress after applying treatments, and 22% said it was physically painful to apply the treatments. Further, adherence to therapy and medication safety were also presented as issues with treatment, particularly for children with AD. Cost of treatment was another concern for some patients, as some reported the substantial cost as a barrier to beginning treatment or continuing treatment. Costs associated with medications and other symptomatic treatment or preventive measures was also reported to vary widely in price.

4. Expectations About the Drug Being Reviewed

Patients living with AD reported the desire for a treatment that eliminates pain and itch, and allows them to pursue a normal life with minimal complications caused by the disease; ultimately, a better quality of life for both patients and their caregivers. One patient commented, “I hope that one day, we can have a solution so people with AD don't have to live in constant pain.” The ESC survey reported that the primary outcome expected by patients is better control of disease that reduces the number of disease flares. The CSPA group goes further, expressing concerns about the use of corticosteroids and the impact on patients’ social/personal lives. The latter is particularly relevant among younger patients, with comments from caregivers regarding severe infection requiring hospitalization, interference with sleep, activities, and a delay of motor skills, as well as embarrassment and self-consciousness due to the condition.

Only the CSPA group reported on patient experience with crisaborole, but this was based on treatments in the US. Regardless, the experience was described as similar to other treatments that had been used by patients in terms of effectiveness and ease of use (the latter was compared with other topical treatments). This includes mixed feedback about the efficacy of the drug, which is consistent with the patient input about the state of current therapies. For example, one patient stated it “didn’t help at all and only caused burning pain when applied,” while another reported that “my daughter has been using eucrisa for a few months and it’s keeping her symptoms manageable. It’s not making it go away but it’s the best her skin has looked in a long time.” In contrast, a different patient reported, “I was given eucrisa to try from a coworker on a bad flare up and I couldn’t believe how fast it

worked and helped !! I went to my dermatologist the following week and got A script for myself I love it! My only complaint is how expensive!" It is important to note again that this survey was conducted in the US when considering the comment about costs. A large proportion of patients (83%) reported experiencing pain, burning, or stinging with the application of the medication; however, those who found it worked well for them reported they were willing to put up with the discomfort.

5. Additional Information

Not applicable.

Appendix 2: Literature Search Strategy

OVERVIEW	
Interface:	Ovid
Databases:	Embase 1974 to present MEDLINE ALL 1946 to present **Note:** Subject headings have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	June 20, 2018
Alerts:	Bi-weekly search updates until October 17, 2018
Study Types:	No search filters were applied
Limits:	No date or language limits were used Conference abstracts were excluded

SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
*	Before a word, indicates that the marked subject heading is a primary topic; or, after a word, a truncation symbol (wildcard) to retrieve plurals or varying endings
.ti	Title
.ab	Abstract
.dq	Candidate term word (Embase)
.ot	Original title
.hw	Heading word; usually includes subject headings and controlled vocabulary
.kf	Author keyword heading word (MEDLINE)
.kw	Author keyword (Embase)
.nm	Name of substance word
.pt	Publication type
.rn	Case Registry/EC number/Name of substance
medall	Ovid database code; MEDLINE ALL (1946–)
oemezd	Ovid database code; Embase 1974 to present, updated daily

MULTI-DATABASE STRATEGY	
Line#	**Search String**
1	(Eucrisa* or crisaborole* or Q2R47HGR7P or PF-06930164 or PF06930164 or AN2728 or AN 2728).ti,ot,ab,kf,rn,hw,nm.
2	1 use medall
3	*crisaborole/ or (Eucrisa* or crisaborole* or PF-06930164 or PF06930164 or AN 2728 or AN2728).ti,ab,kw,dq.
4	3 use oemezd
5	4 not conference abstract.pt.
6	2 or 5
7	remove duplicates from 6

OTHER DATABASES		
PubMed	A limited PubMed search was performed to capture records not found in MEDLINE. Same MeSH, keywords, limits, and study types used as per MEDLINE search, with appropriate syntax used.	
Trial registries (Clinicaltrials.gov and others)	Same keywords, limits used as per MEDLINE search.	

Grey Literature

Dates for Search:	June 2018
Keywords:	Eucrisa, crisaborole, atopic dermatitis
Limits:	No date or language limits used

Relevant websites from the following sections of the CADTH grey literature checklist *Grey Matters: a practical tool for searching health-related grey literature* (https://www.cadth.ca/grey-matters) were searched:

- health technology assessment agencies
- health economics
- clinical practice guidelines
- drug and device regulatory approvals
- advisories and warnings
- drug class reviews
- databases (free)
- internet search.

Appendix 3: Excluded Studies

Table 9: Excluded Studies

Reference	Reason for Exclusion
Eichenfield J Am Acad Derm 2017	No control group (open-label extension)

Appendix 4: Detailed Outcome Data

Table 10: Detailed Efficacy Outcomes

Efficacy Outcomes	AD-301		AD-302	
Exploratory End Points	**Crisaborole N = 503**	**Vehicle N = 256**	**Crisaborole N = 513**	**Vehicle N = 250**
Summary of Signs of Atopic Dermatitis				
Erythema (Redness)				
Baseline 0 (none)	9 (1.8)	5 (2.0)	11 (2.1)	4 (1.6)
1 (mild)	177 (35.2)	88 (34.4)	165 (32.2)	94 (37.6)
2 (moderate)	296 (58.8)	156 (60.9)	315 (61.4)	142 (56.8)
3 (severe)	21 (4.2)	7 (2.7)	22 (4.3)	10 (4.0)
Day 29 0 (none)	142 (29.7)	39 (17.1)	129 (26.5)	32 (14.3)
1 (mild)	248 (51.9)	115 (50.4)	227 (46.7)	92 (41.1)
2 (moderate)	78 (16.3)	66 (28.9)	119 (24.5)	91 (40.6)
3 (severe)	10 (2.1)	8 (3.5)	11 (2.3)	9 (4.0)
Induration/Papulation				
Baseline 0 (none)	6 (1.2)	1 (0.4)	9 (1.8)	3 (1.2)
1 (mild)	139 (27.6)	58 (22.7)	144 (28.1)	57 (22.8)
2 (moderate)	320 (63.6)	173 (67.6)	320 (62.4)	172 (68.8)
3 (severe)	38 (7.6)	24 (9.4)	40 (7.8)	18 (7.2)
Day 29 0 (none)	122 (25.5)	46 (20.2)	114 (23.5)	29 (12.9)
1 (mild)	237 (49.6)	111 (48.7)	225 (46.3)	83 (37.1)
2 (moderate)	104 (21.8)	62 (27.2)	131 (27.0)	102 (45.5)
3 (severe)	15 (3.1)	9 (3.9)	16 (3.3)	10 (4.5)
Exudation (Oozing or Crusting)				
Baseline 0 (none)	226 (44.9)	124 (48.4)	243 (47.4)	141 (56.4)
1 (mild)	162 (32.2)	68 (26.6)	173 (33.7)	61 (24.4)
2 (moderate)	110 (21.9)	60 (23.4)	94 (18.3)	46 (18.4)
3 (severe)	5 (1.0)	1 (0.2)	3 (0.6)	2 (0.8)
Day 29 0 (none)	370 (77.4)	151 (66.2)	362 (74.5)	146 (65.2)
1 (mild)	71 (14.9)	47 (20.6)	80 (16.5)	50 (22.3)
2 (moderate)	34 (7.1)	26 (11.4)	42 (8.6)	23 (10.3)
3 (severe)	3 (0.6)	4 (1.8)	2 (0.4)	5 (2.2)
Excoriation (Evidence of Scratching)				
Baseline 0 (none)	40 (8.0)	22 (8.6)	39 (7.6)	17 (6.8)
1 (mild)	221 (43.9)	88 (34.4)	213 (41.5)	101 (40.4)
2 (moderate)	221 (43.9)	135 (52.7)	233 (45.4)	122 (45.4)
3 (severe)	21 (4.2)	11 (4.3)	28 (5.5)	10 (4.0)
Day 29 0 (none)	253 (52.9)	86 (37.7)	219 (45.1)	76 (33.9)
1 (mild)	145 (30.3)	91 (39.9)	162 (33.3)	75 (33.5)
2 (moderate)	70 (14.6)	44 (19.3)	93 (19.1)	64 (28.6)
3 (severe)	10 (2.1)	7 (3.1)	12 (2.5)	9 (4.0)
Lichenification (Epidermal Thickening)				
Baseline 0 (none)	42 (8.3)	15 (5.9)	28 (5.5)	26 (10.4)
1 (mild)	187 (37.2)	107 (41.8)	215 (41.9)	88 (35.2)
2 (moderate)	244 (48.5)	123 (48.0)	243 (47.4)	122 (48.8)

Efficacy Outcomes	AD-301		AD-302	
3 (severe)	30 (6.0)	11 (4.3)	27 (5.3)	14 (5.6)
Day 29 0 (none)	180 (37.7)	71 (31.1)	168 (34.6)	54 (24.1)
1 (mild)	194 (40.6)	102 (44.7)	197 (40.5)	93 (41.5)
2 (moderate)	93 (19.5)	49 (21.5)	114 (23.5)	68 (30.4)
3 (severe)	11 (2.3)	6 (2.6)	7 (1.4)	9 (4.0)

Source: Clinical study reports for AD-301[5] and AD-302.[6]

Table 11: Subgroup Analyses

Efficacy Outcomes	AD-301		AD-302	
Primary End Points	**Crisaborole N = 503**	**Vehicle N = 256**	**Crisaborole N = 513**	**Vehicle N = 250**
Proportion of Patients Achieving Success in ISGA at Day 29				
By Baseline Age				
2 to 11 years				
Success	32.5%	28.8%	34.3%	15.7%
Failure	67.5%	71.2%	65.7%	84.3%
12 to 17 years				
Success	34.2%	19.2%	26.3%	19.7%
Failure	65.8%	80.8%	73.7%	80.3%
≥ 18 years				
Success	31.6%	22.7%	28.1%	27.7%
Failure	68.4%	77.3%	71.9%	72.3%
By Baseline ISGA (Post Hoc Analysis, Pooled, Both Studies)				
Mild				
Success (pooled)	24.9% (20.46, 29.28)	21.2% (15.08, 27.40)		
Difference (95% CI)	3.6 (–3.94 to 11.21), *P* = 0.3470			
Moderate				
Success (pooled)	36.7% (32.70, 40.69)	22.3% (17.31, 27.29)		
Difference (95% CI)	14.4 (8.01 to 20.79), *P* < 0.0001			

CI = confidence interval; ISGA = Investigator's Static Global Assessment.

Source: Clinical study reports for AD-301[5] and AD-302;[6] manufacturer's response to request for subgroup data.[26]

Appendix 5: Validity of Outcome Measures

Aim

To summarize the validity of the following end point measures:

- Investigator's Static Global Assessment (ISGA)
- Children's Dermatology Life Quality Index (CDLQI)
- Dermatology Life Quality Index (DLQI)
- Dermatitis Family Impact (DFI)
- Severity of Pruritus Scale (SPS).

Findings

Table 12: Validity and Minimal Clinically Important Difference of Outcome Measures

Instrument	Type	Evidence of Validity	MCID	References
ISGA	A 5-point scale that provides a global clinical assessment of AD by investigator, based on signs and symptoms of the condition	No	Not identified	Futamura et al. 2016[21] Eichenfield et al. 2014[2] FDA[30]
DLQI	A self-reported, 10-item questionnaire used to assess six different aspects related to quality of life that are affected by skin diseases, with questions based on a 4-point Likert scale and one-week recall	Yes	3.3 (SRM = 0.27; ES = 0.21)[a]	Finlay and Khan 1994[22] Basra et al. 2008[23] Basra et al. 2015[31]
CDLQI	A self-reported, 10-item questionnaire specifically for children (age 3 to 16 years) used to assess six different aspects related to quality of life that are affected by skin diseases, with questions based on a 4-point Likert scale, based on a one-week recall	Limited	Not identified	Lewis-Jones and Finlay 1995[24]
DFI	An AD-specific, self-administered, 10-item questionnaire designed to assess the impact of disease on the QoL of parents and families of children affected by AD based on a one-week recall	Limited	Unknown	Dodington et al. 2013[25]
SPS	A patient-reported, 4-point rating scale designed to assess pruritus in patients with AD, based on a 24-hour recall	Yes	0.20 (95% CI, 0.18 to 0.22)	Yosipovitch, et al. 2018[32]

AD = atopic dermatitis; CDLQI = Children's Dermatology Life Quality Index; CI = confidence interval; DFI = Dermatitis Family Impact; DLQI = Dermatology Life Quality Index; ES = effect size; ISGA = Investigator's Static Global Assessment; MCID = minimal clinically important difference; QoL = quality of life; SPS = Severity of Pruritus Scale; SRM = standardized response mean.

[a] This is a dermatology-related MCID, which includes but is not specific to AD.

Investigator's Static Global Assessment

The ISGA is a 5-point scale that provides a global clinical assessment of AD severity based on an ordinal scale, scored by an investigator or physician.[21] It is synonymous to the Investigator's Global Assessment (IGA) with the exception of the assessment being static, meaning that the evaluation is done without reference to any other time point.[21] A review of the use of a global assessment in trials of AD by Futamura et al., (2016) reported that the number of points included in the ISGA scale may vary from four to seven, with a 6-point scale ranging from "clear" to "very severe" being used more frequently.[21] The studies of interest for this review use a 5-point scale that ranges from 0 (clear) to 4 (severe). By comparison, the 6-point scale ranges from 0 (clear) to 5 (very severe).[5,6] In the 5-point scale a score of 0 corresponds to a grade of clear, 1 is almost clear, 2 is mild, 3 is moderate, and 4 is severe AD.[5,6] A decrease in score relates to an improvement in signs and symptoms. It was indicated in a guidelines document for the management of AD that the ISGA was designed and commonly used for clinical trials and rarely used in clinical practice.[2] A review of the literature found no information on the validity of the IGA scale specific to patients with AD. Similarly, no information was found regarding what would constitute a minimal clinically important difference (MCID) in patients with AD. The FDA has recommended the IGA (or ISGA) be used as a primary end point for new drugs in AD trials in published draft guidance documents.[30] The ISGA is also widely accepted as the standard outcome used to assess AD in clinical trials according to the clinical expert on this review.

Dermatology Life Quality Index

The DLQI is a widely used dermatology-specific quality of life (QoL) instrument. It is a self-reported, 10-item questionnaire that refers to the preceding week and assesses six different aspects that may affect QoL as a result of living with a dermatological condition.[22,23] The aspects included in the questionnaire are symptoms and feelings, daily activities, leisure, work, or school, personal relationships, and side effects of treatment.[22,23] Each item is scored using a Likert scale that ranges from 0 to 3.[22,23] A score of 0, 1, 2, and 3 corresponds to the following descriptions of how much an aspect is affected by the disease, respectively: "not at all/not relevant," "a little," "a lot," and "very much."[23] The scores of each of the 10 items are summed for an overall DLQI score between 0 and 30 (or a percentage of 30).[22] The higher the score, the greater the impairment of QoL.

The DLQI has shown good test-retest reliability in a population of patients with a variety of skin diseases, 13 (6.5%) of which were diagnosed with atopic eczema.[22] This was based on a Spearman rank correlation between overall DLQI scores of 0.99, $P < 0.0001$ and of individual question scores, which was 0.95 to 0.98, $P < 0.001$.[22] It also demonstrated internal consistency reliability (with Cronbach's alpha coefficients ranging from 0.75 to 0.92 when assessed in 12 international studies) and construct validity (as 37 separate studies have mentioned a significant correlation of the DLQI with either generic or dermatology-specific and disease-specific measures); however, the studies were also based on populations that were not specific to AD.[23] The responsiveness of the DLQI was also assessed, and the DLQI was determined to be capable of detecting changes before and after treatment in patients with AD in 17 different studies.[31]

A study by Basra, et al. (2015[31]) used an anchor-based approach to assess the responsiveness of the DLQI and determine the MCID. Patients older than 16 years of age without significant comorbidities were recruited from dermatology outpatient clinics following a first-time referral or follow-up patient for various inflammatory skin conditions.[31] A total of

192 patients suffering from different skin conditions ranging from acute to chronic were included (50.5% with psoriasis, 21.9% with acne, 12.5% with eczema, and 15.1% other).[31] Detail regarding the disease severity of patients was not reported. The Global Question (GQ) score was based on a self-assessment of the patient's skin disease using a visual analogue scale from 0 to 10 (from clear to worst possible). The mean GQ score for the inclusion patients was 5.7 (SD = 2.5). They were asked to complete the DLQI and answer the GQ as an initial assessment, and again between one and three months later. The follow-up also included the completion of a Global Rating of Change Questionnaire (GRCQ) to assess the overall change in QoL.[31] Patients were divided into four categories relating to the magnitude of change in overall QoL as recorded by the results of the GRCQ. A "small change," indicated by a score of plus or minus 2 or 3 corresponded to a mean change in DLQI of 3.3 (standardized response mean = 0.27; effect size = 0.21), which was suggested as the MCID for the DLQI.[31]

There are a few limitations with the determination of the MCID for the DLQI. The anchor-based approach was based on a subjective, global assessment of change that is subject to recall bias, and it is not specific to one skin condition or AD in particular.[31] The authors did address this and noted that the use of a global assessment was reasonable because the population was comprised of patients with a mix of diagnosed skin conditions; however, the lack of specificity to AD is another limitation and should be taken into consideration when applied to an AD-specific population.

Children's Dermatology Life Quality Index

The CDLQI was developed using the same methods for the development of the DLQI, in children between the ages of 3 years and 16 years.[24] Like the DLQI, the CDLQI is a self-reported questionnaire that refers to the preceding week, used to assess the impact of skin disease on the QoL but for children and may be completed with help from a parent/guardian. It also involves 10 questions that address the following topics: symptoms and feelings, leisure, school or holidays, personal relationships, sleep, and treatment. A score of 0,1, 2, or 3 is assigned, respectively, to the following answers to each of the 10 questions: "not at all," "only a little," "quite a lot," or "very much."[24]

Test-retest reliability was assessed in a group of 46 patients with stable conditions.[24] Patients were asked to complete the CDLQI twice, four days apart. The mean difference between pairs was 0.33 (SD = 2.5, *P* was not significant), and the Spearman's correlation coefficient was 0.86, which indicated acceptable reliability according to the authors of the study.[24] Like the other instruments discussed, this was conducted in a group of patients with mixed dermatological conditions and is therefore not specific to AD. Further, other information relating to the validity of the CDLQI or an MCID was not identified, which was also reported by a systematic review of QoL instruments for children with eczema conducted in 2017.[33]

Dermatitis Family Impact

The DFI questionnaire was designed to assess the impact of disease on the QoL of parents and families of children affected by AD.[25] It is a disease-specific, self-administered questionnaire that relies on a one-week recall, and consists of 10 items that were derived from ethnographical interviews and focus groups.[25] The 10 items of the questionnaire address the following topics: housework, food preparation, sleep, family leisure activity, shopping, expenditure, tiredness, emotional distress, relationships, and the impact on the carer's life due to helping with treatment.[25,34] Each question is scored on a four-point Likert

scale ranging from 0 to 3, for an overall total ranging from 0 to 30, with a higher score corresponding to a greater negative impact on the family's QoL due to AD.[25] Dodington, et al. (2013) conducted a review of the measurement properties of the DFI and reported evidence of internal consistency in three non-English versions of the questionnaire, and test-retest reliability. The latter was based on a study conducted in the UK[35] that reported a Spearman rank correlation coefficient of 0.95; however, this was specific to infants, or children under the age of four years. The review also reported evidence of convergent validity demonstrated by a correlation of the DFI to other dermatology-specific QoL instruments in 26 studies; however, none of these studies were conducted in North America. In summary, while there is evidence of validation for the DFI, the applicability to a Canadian or even North American context is questionable. This is relevant as the DFI is a subjective measure of the family's perspective and how they are impacted by AD; which may vary due to cultural considerations. Recall bias is also a limitation of the DFI, which is based on a one-week recall that should represent the perspective of more than one person.[25] Lastly, an MCID was not identified for the DFI.

Severity of Pruritus Scale

Studies AD-301 and AD-302 used a SPS to assess the extent or severity of itching in patients. It is scored on a four-point numeric rating scale that ranges from 0 to 3. The ratings correspond to a grade and definition, where 0 is a grade of none or "no itching;" 1 is mild or "occasional, slight itching/scratching;" 2 is moderate or "constant or intermittent itching/scratching which is not disturbing sleep;" and 3 is severe or "bothersome itching/scratching which is disturbing sleep." Following the provided instructions, the scale was completed by study participant or parent/guardian using an electronic diary and based on a 24 hour recall.[5,6]

An assessment of the psychometric properties of the SPS was provided in a published report supported by the manufacturer, based on data from the phase II and phase III studies for crisaborole.[32] The intention-to-treat population (n = 1,522) with baseline ISGA data were used to conduct the assessment. The intraclass correlation coefficient was used to support test-retest reliability of the SPS, which was calculated using data from patients who did not experience a change in ISGA score during between day 1 and day 8 of the study.[32] The number of patients included in this subset was not reported; however, they noted an intraclass correlation of 0.70 was observed when an average of two SPS observations were used, which just meets the cut-off for acceptable test-retest reliability.[32] Pearson's correlation coefficient was also calculated for the SPS and the ISGA, DLQI, CDLQI, DFI, and signs of AD, in an effort to examine the convergent validity of the SPS. A moderate correlation (ranging from 0.50 to 0.59) was reported for each of the outcomes, with the exception of the signs of AD, which were all below 0.50.[32] The study also assessed the ability of the SPS to detect change by plotting the SPS data as a function of the ISGA data (both as a categorical and continuous variable) and an assessment using a repeated-measures longitudinal mixed model approach. The results were only descriptive, highlighting that the SPS changed accordingly with a change in disease severity indicated by the ISGA.[32] Finally, an MCID of 0.20 (95% confidence interval of 0.18 to 0.22) was reported using a repeated-measures longitudinal model linked to a difference of one ISGA category, as an anchor.

The information presented for the validation of the SPS should be carefully considered as some of the methodology used was only described briefly, sample sizes were not explicitly reported, and analyses were done post hoc. Further, the SPS was validated using the data

it is intended to describe, in an assessment sponsored by the manufacturer, both of which are limitations to be mindful of.

Conclusions

The ISGA is one of the most commonly used tools in clinical trials to evaluate disease severity in patients with AD, despite limited published evidence of validation. However, the ISGA is recommended by the FDA as the primary end point to use in the evaluation of new drugs indicated for AD. The DLQI is also a frequently used tool, but to assess the HRQoL in patients with AD. Evidence of validation, including good test-retest reliability, internal consistency reliability, and construct validity are available for the DLQI, but the assessment of validation was not specific to patients with AD. An MCID of 3.3 specific to patients with AD for the DLQI was also provided. A version of the DLQI was developed for children, i.e., the CDLQI, which is commonly used despite a lack of evidence for validity. The DFI was specifically designed to assess the impact of AD on a patient's family or carers, which is widely used internationally and has been translated for a variety of languages. While there is validation of the DFI in non-English versions or studies conducted outside of North America, limited evidence of validity was provided for the North American context, and no evidence was identified regarding an MCID. Finally, the SPS was developed to assess pruritus in patients with AD, with evidence of validity provided by a study sponsored by the manufacturer. The quality of evidence is poor; however, an MCID of 0.20 was reported.

Appendix 6: Summary of Other Studies

Objective

To summarize the results of the long-term safety extension (LTSE; AD-303) that evaluated the long-term safety of crisaborole topical ointment, 2% in patients at least two years of age with mild-to-moderate atopic dermatitis (AD).

Findings

Study Design

The study design characteristics of the multi-centre, open-label, LTSE study are summarized in Table 13. Patients of at least two years of age with mild-to-moderate AD (based on an ISGA score of 2 or 3) were eligible to participate in the LTSE if they had completed the pivotal studies (AD-301 and AD-302) without experiencing any crisaborole-related adverse event (AE) or serious adverse event (SAE) that would prevent them from further treatment with the drug (N = 517). The safety assessment included data from the pivotal studies, which were four weeks long, in addition to the 48-week LTSE, for a total study duration of 52 weeks. The 52 weeks was broken down into 12, four-week (28-day) treatment cycles. At the beginning of each treatment cycle, i.e., every 28 days, the patient was assessed using the Investigator's Static Global Assessment (ISGA) to determine whether they would proceed to the next treatment cycle according to an "on-treatment" protocol or "off-treatment" protocol. An ISGA score of ≥ 2 warranted the former, which involved initiation or continuation of the intervention, which was to apply the crisaborole ointment twice daily for 28 days. During the "off-treatment" protocol, treatment with crisaborole was not initiated, but patients were permitted to use non-medicated emollients as needed. The use of corticosteroid or calcineurin inhibitors was prohibited, unless prescribed by the principal investigator or a designee, and could not be used simultaneously with the study drug. Sunbathing, tanning bed use, or light therapy (e.g., ultraviolet, ultraviolet-B, or psoralen plus ultraviolet-A) were also prohibited.

Table 13: Study Design and Characteristics

		AD-303
Design and Population	**Study design**	Multi-centre, open-label, long-term safety study
	Participants (N)	517
	Eligibility	Patients age ≥ 2 years old with mild-to-moderate AD (ISGA score of 2 or 3) who completed the pivotal studies (AD-301, AD-302) without experiencing crisaborole treatment-related AE or SAE that precluded further treatment with crisaborole ointment
	Primary objective	To assess the long-term safety of crisaborole ointment
Drug	**Intervention**	Crisaborole topical ointment, 2% applied twice daily for 28 days to all treatable AD-involved areas (excluding the scalp) during on-treatment periods; i.e., when the patient's ISGA score was ≥ 2 during the patient assessment Off-treatment periods were initiated when a patients ISGA score was clear (0) or almost clear (1)
	Comparators	None
Duration	**Pivotal study treatment**	4 weeks
	Safety extension treatment	12 treatment cycles (28 days or 4 weeks per cycle)
Outcomes	**Primary end points**	Treatment-emergent AEs Treatment-emergent SAEs

AD = atopic dermatitis; AE = adverse event; ISGA = Investigator's Static Global Assessment; SAE = serious adverse event.

Reference: Clinical study report.[29]

Methods

The primary objective of the study was to evaluate the long-term safety of crisaborole topical ointment, 2% in patients aged two years or older with mild-to-moderate AD. The primary end points were treatment-emergent adverse events and treatment-emergent serious adverse events, referred to as AEs and SAEs herein. AEs were defined as events that occurred on or after the day consent/assent was provided up until the end of the LTSE (study day 337). The frequency of AEs was reported by age group (age 2 to 11 years; age 12 to 17 years; and age ≥18 years), and by 12-week periods (days 1 to 85; days 86 to 169; days 170 to 253; and day 254). The safety assessment and results reported by age group included AEs that occurred during the pivotal trials as well (i.e., AD-301, AD-302, and AD-303), but the results reported by 12-week periods only include AEs that occurred after the initial four weeks or pivotal trials (i.e., exclusive to AD-303). Results were presented descriptively; no statistical analyses were performed.

Patient Disposition

The patient disposition is summarized in Table 14. A total of 517 (34.0%) patients randomized to the preceding double-blind, pivotal trials proceeded to the open-label extension study and were enrolled in the LTSE. Of those who were enrolled, 357 (69%) had previously received treatment with crisaborole, and 160 (31%) had received vehicle (Table 4). A total of 271 (52.4%) patients completed the safety study, with a total number of 246 (47.6%) of patients having discontinued the study. The highest discontinuation rate was among the 2 to 11 year age group (50.6%), followed by 45.9% in the 12 to 17 year age group and 36.5% in the greater than and equal to 18 years age group. The most common

reason for early discontinuation (aside from "other," 22.2% overall) was withdrawal by parent or guardian (12.2% overall, or 15.6%, 10.3%, and 0% for the 2 to 11 years, 12 years to 17 years, and ≥18 years age groups, respectively), followed by lost to follow-up (7.0% overall), withdrawal by patient (4.4%), and AEs (1.7%). No deaths were reported in this study.

Table 14: Patient Disposition for AD-303

Disposition, n (%)	Age 2 to 11 y	Age 12 to 17 y	Age ≥ 18 y	Total
Study Status				
Enrolled	308 (59.6)	146 (28.2)	63 (12.2)	517 (100)
Completed	152 (49.4)	79 (54.1)	40 (63.5)	271 (52.4)
Discontinued	156 (50.6)	67 (45.9)	23 (36.5)	246 (47.6)
Reason for Early Discontinuation				
AEs	5 (1.6)	1 (0.7)	3 (4.8)	9 (1.7)
Withdrawal by patient	4 (1.3)	12 (8.2)	7 (11.1)	23 (4.4)
Withdrawal by parent or guardian	48 (15.6)	15 (10.3)	0	63 (12.2)
Lost to follow-up	25 (8.1)	6 (4.1)	5 (7.9)	36 (7.0)
Death	0	0	0	0
Other	74 (24.0)	33 (22.6)	8 (12.7)	115 (22.2)

AE = adverse event; y = years.

Source: Clinical study report.[29]

Baseline Characteristics

The baseline characteristics that were reported for the population included in the LTSE are summarized in Table 15.Patients had a mean age of 11.7 years (standard deviation [SD] 10.39), and 56.5% and 43.5% of the population was 2 to 6 years and 7 to 11 years old, respectively. Approximately 59% of the patients were female and 41% were male. Further, the majority of patients were white (60.9%) or black/African-American (29.4%). The rest of the population identified as Asian (5.4%), American Indian/Alaska Native (0.2%), or other (3.9%).

Table 15: Baseline Characteristics in AD-303

Characteristics	Age 2 to 11 y (N = 308)	Age 12 to 17 y (N = 146)	Age ≥ 18 y (N = 63)	Total (N = 517)
Age, mean (SD)	6.1 (2.84)	14.0 (1.49)	34.0 (13.38)	11.7 (10.39)
Age subcategory, n (%)				
2 to 6 y	174 (56.5)	NA	NA	174 (56.5)
7 to 11 y	134 (43.5)	NA	NA	134 (43.5)
Sex, n (%)				
Male	131 (42.5)	61 (41.8)	19 (30.2)	211 (40.8)
Female	177 (57.5)	85 (58.2)	44 (69.8)	306 (59.2)
Ethnicity				
Hispanic/Latino	47 (15.3)	26 (17.8)	9 (14.3)	82 (15.9)
Not Hispanic/Latino	261 (84.7)	120 (82.2)	54 (85.7)	435 (84.1)
Race, n (%)				
American Indian or Alaska Native	1 (0.3)	0	0	1 (0.2)

Characteristics	Age 2 to 11 y (N = 308)	Age 12 to 17 y (N = 146)	Age ≥ 18 y (N = 63)	Total (N = 517)
Asian	17 (5.5)	6 (4.1)	5 (7.9)	28 (5.4)
Black or African-American	84 (27.3)	44 (30.1)	24 (38.1)	152 (29.4)
Native Hawaiian or Other Pacific Islander	1 (0.3)	0	0	1 (0.2)
White	189 (61.4)	94 (64.4)	32 (50.8)	315 (60.9)
Other	16 (5.2)	2 (1.4)	2 (3.2)	20 (3.9)

SD = standard deviation; y = years.

Source: Clinical study report.[29]

A summary of exposure to study treatment during the main pivotal trials (AD-301 and AD-302) in addition to the LTSE (AD-303), has been provided in Table 4. When broken down by 12-week periods, 93.2% of the total number of patients enrolled was included in the week 1 to 12 period, 82.8% in weeks 13 to 24, 71.2% in weeks 25 to 36, and 43.7% from weeks 36 to 48. The number of patients included in each 12-week period was also broken down by age group. The proportion of the total number of patients per age group was similar across groups for the first half of the LTSE study up to week 24, but varied by age group after that point. During weeks 25 to 36, and weeks 36 to 48, 69.5% and 40.3% of the 2 to 11 years age groups, respectively, were included. Regarding the 12 to 17 years age group, 71.9% and 45.2% of patients were included at weeks 25 to 36 and weeks 36 to 48, respectively. The adult group (greater than and equal to 18 years of age) included 77.8% of patients during week 25 to 36 and 57.1% at weeks 36 to 48. The mean number of applications of the study drug was similar across age groups, with an overall mean of 348.9 (SD = 183.30). The mean amount of drug used was highest among the 2 to 11 years age group at 793.46 g (SD = 1,039.69), followed by the 12 to 17 years age group at 791.13 g (SD = 1,052.15) and 528.32 g (SD = 722.22) among the adult age group. The amount of drug used per application was also reported, and was similar across age groups (ranging from 2.10 g in the adult group to 2.40 g in the youngest group) for an overall mean of 2.34 g (SD = 2.55). Lastly, the number of on-treatment periods and duration of on-treatment periods were also similar across age groups, with an overall mean of 6.2 treatment periods (SD = 3.20) and 28.4 days (SD = 6.10), respectively.

Table 16: Summary of Treatment Exposure

Treatment Exposure	Age 2 to 11 y (N = 308)	Age 12 to 17 y (N = 146)	Age ≥ 18 y (N = 63)	Total (N = 517)
Treatment Received in AD-301 or AD-302, n (%)				
Crisaborole ointment	216 (70.1)	102 (69.9)	39 (61.9)	357 (69.1)
Vehicle	92 (29.9)	44 (30.1)	24 (38.1)	160 (30.9)
Patients Included in Each 12-Week Period, n (%)				
Week 1 to 12	288 (93.5)	136 (93.2)	58 (92.1)	482 (93.2)
Week 13 to 24	254 (82.5)	121 (82.9)	53 (84.1)	428 (82.8)
Week 25 to 36	214 (69.5)	105 (71.9)	49 (77.8)	368 (71.2)
Week 36 to 48	124 (40.3)	66 (45.2)	36 (57.1)	226 (43.7)
Number of Applications[a]				
Patients, n	304	146	60	510
Mean (SD)	349.0 (179.57)	349.4 (193.21)	347.7 (180.33)	348.9 (183.30)
Amount of Drug Used,[a,b] g				
Patients, n	308	146	63	517

Treatment Exposure	Age 2 to 11 y (N = 308)	Age 12 to 17 y (N = 146)	Age ≥ 18 y (N = 63)	Total (N = 517)
Mean (SD)	793.46 (1,039.62)	791.13 (1,052.15)	528.32 (722.22)	760.49 (1,012.07)
Amount of Drug Used per Application,[a] g				
Patients, n	304	146	60	510
Mean (SD)	2.40 (2.50)	2.29 (2.38)	2.10 (3.20)	2.34 (2.55)
Number of On-Treatment Periods,[a] n				
Patients, n	308	146	63	517
Mean (SD)	6.2 (3.14)	6.3 (3.35)	5.9 (3.12)	6.2 (3.20)
Duration of On-Treatment Periods,[a] Days				
On-treatment periods, n	1,903	921	370	3,194
Mean (SD), days	28.4 (5.83)	28.3 (6.52)	28.6 (6.40)	28.4 (6.10)

SD = standard deviation; y = years.

[a] Includes data from both the main pivotal trials (AD-301 and AD-302) and the LTSE (AD-303).

[b] Amount of drug used was set to 0 for patients who were not dispensed crisaborole ointment and for patients with no complete dispense and return weight records.

Reference: Clinical study report.[29]

Results

Nearly two-thirds (65%) of patients experienced at least one AE (Table 17). The most frequently reported AE was dermatitis atopic (11.2%), which was reported upon worsening, exacerbation, flare, or flare-up of an existing condition.[36] The next most frequently reported AE reported in greater than and equal to 2% of patients was upper respiratory tract infection (URTI) (10.3%), followed by nasopharyngitis (7.7%), cough (6.8%), pyrexia (5.6%), sinusitis (4.8%), pharyngitis streptococcal (3.9%), oropharyngeal pain (3.7%), application site infection (3.5%), asthma (3.1%), vomiting (2.9%), eczema (2.5%), pharyngitis (2.3%), influenza (2.3%), ear infection (2.3%), application site pain (2.3%), diarrhea (2.3%), headache (2.1%), viral infection(2.1%), otitis media(2.1%), and seasonal allergy (2.1%). Only 10.3% of subjects reported at least one AE that was at least possibly related to the study drug.

SAEs were reported in 1.7% of patients, all of whom were under the age of 18 (Table 17). Each SAE was only reported once, i.e., in one (0.2%) of patients. The following were reported as having occurred in patients between the age of 2 years and 11 years: appendicitis, application site infection, URTI, laceration, central nervous system (CNS), ventriculitis, and asthma. Anaphylactic reaction, depression, and suicide attempt were reported in the subgroup of patients between the age of 12 years and 17 years. No deaths were reported in this LTSE and none of the reported SAEs were believed to be related to the study drug.

Table 17: Summary of Adverse Events by Age Group

Harms, n (%)	Age 2 to 11 y (N = 308)	Age 12 to 17 y (N = 146)	Age ≥ 18 y (N = 63)	Total (N = 517)
Subjects reporting ≥ 1 AE	209 (67.9)	95 (65.1)	32 (50.8)	336 (65.0)
Subjects reporting ≥ 1 SAE	6 (1.9)	3 (2.1)	0	9 (1.7)
Deaths	0	0	0	0
AEs Reported in ≥ 2% of Patients				
Gastrointestinal disorders	31 (10.1)	8 (5.5)	5 (7.9)	44 (8.5)
Diarrhea	10 (3.2)	1 (0.7)	1 (1.6)	12 (2.3)
Vomiting	14 (4.5)	0	1 (1.6)	15 (2.9)
General disorders and administration site conditions	41 (13.3)	12 (8.2)	5 (7.9)	58 (11.2)
Application site pain	6 (1.9)	5 (3.4)	1 (1.6)	12 (2.3)
Pyrexia	27 (8.8)	2 (1.4)	0	29 (5.6)
Immune system disorders	11 (3.6)	2 (1.4)	2 (3.2)	15 (2.9)
Seasonal allergy	10 (3.2)	1 (0.7)	0	11 (2.1)
Infections and infestations	157 (51.0)	56 (38.4)	14 (22.2)	227 (43.9)
Application site infection	14 (4.5)	4 (2.7)	0	18 (3.5)
Ear infection	8 (2.6)	2 (1.4)	2 (3.2)	12 (2.3)
Influenza	9 (2.9)	1 (0.7)	2 (3.2)	12 (2.3)
Nasopharyngitis	21 (6.8)	15 (10.3)	4 (6.3)	40 (7.7)
Otitis media	11 (3.6)	0	0	11 (2.1)
Pharyngitis	8 (2.6)	4 (2.7)	0	12 (2.3)
Pharyngitis streptococcal	17 (5.5)	3 (2.1)	0	20 (3.9)
Sinusitis	17 (5.5)	6 (4.1)	2 (3.2)	25 (4.8)
URTI	38 (12.3)	12 (8.2)	3 (4.8)	53 (10.3)
Viral infection	6 (1.9)	3 (2.1)	2 (3.2)	11 (2.1)
Nervous system disorders	7 (2.3)	9 (6.2)	2 (3.2)	18 (3.5)
Headache	5 (1.6)	5 (3.4)	1 (1.6)	11 (2.1)
Respiratory, thoracic, and mediastinal disorders	55 (17.9)	26 (17.8)	5 (7.9)	86 (16.6)
Asthma	11 (3.6)	4 (2.7)	1 (1.6)	16 (3.1)
Cough	27 (8.8)	6 (4.1)	2 (3.2)	35 (6.8)
Oropharyngeal pain	11 (3.6)	7 (4.8)	1 (1.6)	19 (3.7)
Skin and subcutaneous tissue disorders	65 (21.1)	35 (24.0)	9 (14.3)	109 (21.1)
Dermatitis atopic	37 (12.0)	16 (11.0)	5 (7.9)	58 (11.2)
Dermatitis contact	8 (2.6)	6 (4.1)	2 (3.2)	16 (3.1)
Eczema	10 (3.2)	2 (1.4)	1 (1.6)	13 (2.5)
SAEs				
Immune system disorders	0	1 (0.7)	0	1 (0.2)
Anaphylactic reaction	0	1 (0.7)	0	1 (0.2)
Infections and infestations	3 (1.0)	0	0	3 (0.6)
Appendicitis	1 (0.3)	0	0	1 (0.2)
Application site infection	1 (0.3)	0	0	1 (0.2)
URTI	1 (0.3)	0	0	1 (0.2)
Injury, poisoning, and procedural complications	1 (0.3)	0	0	1 (0.2)
Laceration	1 (0.3)	0	0	1 (0.2)

Harms, n (%)	Age 2 to 11 y (N = 308)	Age 12 to 17 y (N = 146)	Age ≥ 18 y (N = 63)	Total (N = 517)
Nervous system disorders	1 (0.3)	0	0	1 (0.2)
CNS ventriculitis	1 (0.3)	0	0	1 (0.2)
Psychiatric disorders	0	2 (1.4)	0	2 (0.4)
Depression	0	1 (0.7)	0	1 (0.2)
Suicide attempt	0	1 (0.7)	0	1 (0.2)
Respiratory, thoracic, and mediastinal disorders	1 (0.3)	0	0	1 (0.2)
Asthma	1 (0.3)	0	0	1 (0.2)

AE = adverse event; CNS = central nervous system; SAE = serious adverse event; URTI = upper respiratory tract infection; y = years.

Note: All AEs and SAEs reported were treatment-emergent, unless otherwise specified.

Source: Clinical study report.[29]

Treatment-emergent AEs and SAEs were also reported by 12-week periods, and do not include AEs that occurred prior to the start of the LTSE, i.e., during the pivotal trials (Table 18). Also of note, the AEs by 12-week periods were reported by a subject frequency, rather than event frequency. The number of subjects reporting at least one AE during the first 85 days of the LTSE study was 171 (35.5%), followed by 159 (37.1%) between days 86 and 169, 121 (32.9%) between days 170 to 254, and 73 (32.3%) after 254 days. AEs that were reported in greater than and equal to 1% of patients across groups, or during each time period, included dermatitis atopic, URTI, nasopharyngitis, and pyrexia. Dermatitis atopic was reported in between 4.4% and 5.4% of patients throughout the study, with the exception of between days 170 and 253 (2.4%). Infections and infestations in general remained fairly consistent throughout the LTSE, reported in between 15.9% and 20.8% of patients during each time period. General disorders and administration site conditions were highest during the first time period (4.8%), and then remained between 2.2% and 2.8% during each successive time period. The specific results for pyrexia were similar (ranging between 1.2% and 2.7% throughout).

As for the SAEs by 12-week periods, which are also summarized in Table 18, the majority of subjects reporting an SAE had reported them during the first time period. This included a report of application site infection, URTI, CNS ventriculitis, and asthma, all of which occurred in one patient (0.2%) each. An anaphylactic reaction was the only SAE reported in the second time period (days 86 to 169). Two reports of psychiatric disorders (depression and a suicide attempt) were reported during the third time period (days 170 to 253), without any additional SAEs after that time.

Table 18: Summary of Adverse Events by 12-Week Periods

	Day 1 to 85 (N = 482)	Day 86 to 169 (N = 428)	Day 170 to 253 (N = 368)	Day ≥ 254 (N = 226)
Subjects reporting ≥ 1 AE	171 (35.5)	159 (37.1)	121 (32.9)	73 (32.3)
Subjects reporting ≥ 1 SAE	4 (0.8)	1 (0.2)	2 (0.5)	0
Deaths	0	0	0	0
AEs Reported in ≥ 1% of Patients Across Groups				
General disorders and administration site conditions	23 (4.8)	12 (2.8)	9 (2.4)	5 (2.2)
Pyrexia	13 (2.7)	5 (1.2)	7 (1.9)	3 (1.3)
Infections and infestations	93 (19.3)	89 (20.8)	65 (17.7)	36 (15.9)
Nasopharyngitis	15 (3.1)	12 (2.8)	6 (1.6)	3 (1.3)
URTI	19 (3.9)	16 (3.7)	18 (4.9)	6 (2.7)
Skin and subcutaneous tissue disorders	42 (8.7)	36 (8.4)	13 (3.5)	28 (12.4)
Dermatitis atopic	25 (5.2)	23 (5.4)	9 (2.4)	10 (4.4)
SAEs				
Immune system disorders	0	1 (0.2)	0	0
Anaphylactic reaction	0	1 (0.2)	0	0
Infections and infestations	2 (0.4)	0	0	0
Application site infection	1 (0.2)	0	0	0
URTI	1 (0.2)	0	0	0
Nervous system disorders	1 (0.2)	0	0	0
CNS ventriculitis	1 (0.2)	0	0	0
Psychiatric disorders	0	0	2 (0.5)	0
Depression	0	0	1 (0.3)	0
Suicide attempt	0	0	1 (0.3)	0
Respiratory, thoracic, and mediastinal disorders	1 (0.2)	0	0	0
Asthma	1 (0.2)	0	0	0

AE = adverse event; CNS = central nervous system; SAE = serious adverse event; URTI = upper respiratory tract infection.

Note: All AEs and SAEs reported were treatment-emergent, unless otherwise specified.

Source: Clinical study report.[29]

Limitations

There are a few limitations to note for the long-term, open-label, non-randomized safety extension of the two pivotal trials for crisaborole. Although the results of the safety assessment did not reveal any safety-related signals, it is unclear whether the AEs that were reported were due to a natural course of the disease or were attributed to the use of crisaborole as the study was not controlled. In addition, rescue therapy was permitted at the discretion of the investigator or designee, which may also affect the true assessment of safety for crisaborole. The LTSE study was also an open-label trial design, where investigators and participants are not blinded to treatment allocation, which may have an effect on subjective outcomes like patient-reported AEs. The discontinuation rate of 47.6% should also be noted as a limitation, with about half (49.4%) of patients in the 2- to 11-year-old age group having completed the study. The most common reason for discontinuation, other than "other," was withdrawal by parent or guardian (12.2% overall). Again, there was an absence of safety signals uncovered in this study, but long-term data are not available

for nearly two-thirds of the patients enrolled in the main pivotal trials (34.0% enrolled from AD-301 and AD-302), which was further reduced by a discontinuation rate of 47.6%, which makes this conclusion difficult to assess. However, the low enrolment may in part be due to the eligibility criteria for the LTSE as patients were required to have a diagnosis of mild-to-moderate AD (ISGA of 2 or 3) to continue. A total of 395 (52.0%) and 440 (57.7%) participants from AD-301 and AD-302, respectively, met this criterion based on the secondary efficacy outcome (failure to meet an ISGA score of clear or almost clear at day 29) in the main pivotal trials. Lastly, statistical analyses were not reported for any of the outcomes included, which limits the interpretability of the results.

Summary

In summary, the LTSE (AD-303) of AD-301 and AD-302 did not reveal any new signals regarding the safety of treatment with crisaborole ointment, 2% in patients with mild-to-moderate AD, based on a 48-week assessment, in addition to safety data from the pivotal trials. However, the interpretation of the safety results are limited by the open-label nature of the study, lack of a control group and statistical testing, in addition to a high discontinuation rate in the study. These limitations must be considered along with the interpretation of these results due to the uncertainty they introduce to the evaluation.

Appendix 7: Summary of Indirect Comparisons

Introduction

Background

The clinical trials included in this review did not provide direct evidence regarding the comparative efficacy and safety of crisaborole ointment relative to other topical therapies, such as topical calcineurin inhibitors (TCI) and topical corticosteroids (TCS). The aim of this section is to provide an overview and critical appraisal of the published and unpublished indirect evidence available for the assessment of the comparative efficacy and harms of crisaborole 2% ointment to the available topical pharmacologic therapies in patients with mild-to-moderate atopic dermatitis (AD).

Methods

One network meta-analysis (NMA) by Hughes et al. was included in the manufacturer's pharmacoeconomic evaluation.[37] In addition, CDR conducted an independent literature search for published indirect treatment comparisons that compared crisaborole ointment with other available topical therapies when used for the treatment of mild-to-moderate AD. One additional NMA conducted by the Institute for Clinical and Economic Review (ICER) was identified from the CADTH Common Drug Review (CDR) literature search.[38]

Description of Indirect Treatment Comparisons Identified

The inclusion criteria for each of the NMAs are summarized in Table 19 below.

Table 19: PICOS Criteria for Study Inclusion

	Manufacturer's NMA	ICER's NMA
Population	Children (≥ 2 years) and adults with a clinical diagnosis of mild-to-moderate AD	Moderate-to-severe AD (for dupilumab vs. placebo) Mild-to-moderate AD (for crisaborole vs. emollient)
Interventions	• Crisaborole 2% ointment • TCS • TCI • Other PDE4 inhibitors • Topical agents under evaluation for mild-to-moderate or moderate AD • Vehicle	Dupilumab (FDA-approved dosage[a]) vs. placebo Crisaborole (FDA-approved dosage[a]) vs. emollient
Comparisons	Comparisons were made between the above mentioned regimens	
Outcomes	Efficacy Outcomes (All time points with priority on days 8, 29, 36) • AD severity: ISGA scores (or other AD severity scales) • Proportion with ≥ 2-grade improvement in ISGA (or similar scale) to clear (0) or almost clear (1) • Proportion with clear (0) or almost clear (1) in ISGA (or similar scale) • Other common AD severity scales (in case ISGA is not available)	Clinical benefits, i.e., ISGA, pruritus, HRQoL Harm • Treatment-related AE • Skin infection

	Manufacturer's NMA	ICER's NMA
	• Time to improvement • Severity of signs of AD • Erythema, exudation, excoriation, induration/papulation, lichenification • HRQoL • Safety Outcomes • AEs • SAEs • Any cutaneous AE • Any systemic steroid-related AE • WDAEs	
Study Design	RCTs (parallel and crossover) SRs and MAs of RCTs	RCTs, comparative observational studies, and high-quality SRs

AD = atopic dermatitis; AE = adverse event; HRQoL = health-related quality of life; ICER = incremental cost-effectiveness ratio; MA = meta-analysis; NMA = network meta-analysis; PDE4 = phosphodiesterase type 4; PICOS = Population, Intervention, Comparison, Outcome, Study Type; RCT = randomized controlled trial; SAE = serious adverse event; SR = systematic review; TCI = topical calcineurin inhibitors; TCS = topical corticosteroids; vs. = versus; WDAE = withdrawal due to adverse event.

[a] Details regarding the FDA-approved dosage for dupilumab and crisaborole were not provided in this study.

Source: Manufacturer-provided NMA;[37] ICER report.[38]

Review and Appraisal of Indirect Treatment Comparisons

Review of Manufacturer-Provided Network Meta-Analysis[37]

Objectives and Rationale for Manufacturer's Network Meta-Analysis

The objective of this report was to evaluate the comparative efficacy, safety, and health-related quality of life (HRQoL) impact of crisaborole 2% ointment versus other topical pharmacologic therapies, for the treatment of mild-to-moderate AD, using an NMA approach.

Methods for Manufacturer's Network Meta-Analysis

Study Eligibility and Selection Process

The authors indicated that a systematic review and associated NMA were conducted according to the requirements of major health technology assessment agencies such as the National Institute for Health and Care Excellence (NICE) and the Cochrane Collaboration and reported according to the general guidelines described by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement.

The NMA was based on the systematic review of the literature which included both electronic and manual search components. Multiple databases were searched from inception to September 1, 2017. The search was restricted to articles published in English.

It is unclear whether the selection criteria were defined a priori. The main inclusion criteria for the systematic review were randomized controlled trials (RCTs) that recruited children (two years and older) and adult patients with a clinical diagnosis of mild-to-moderate AD. Only topical therapies were included. To be eligible, the studies were required to report at least one of the following outcomes: change in AD severity, HRQoL, or safety. Study selection was accomplished through two levels of screening by two independent researchers. Any disagreements were resolved through discussion and consensus.

Data Extraction

Data were extracted by one reviewer, and verified by a second reviewer. Any disagreements were resolved by consensus.

Comparators

Topical pharmacological therapies such as TCS and TCI were of interest for inclusion in the NMA.

Outcomes

The main end points of interest included in the systematic review were stated to be:

AD severity: measured with Investigator's Static Global Assessment (ISGA) scores (or other AD severity scales), proportion of patients with greater than and equal to 2-grade improvement in ISGA (or similar scale) to clear (0) or almost clear (1), at the time points of day 7 to 8; day 14 to 15; day 21 to 22; day 28 to 29; and day 42 to 43.

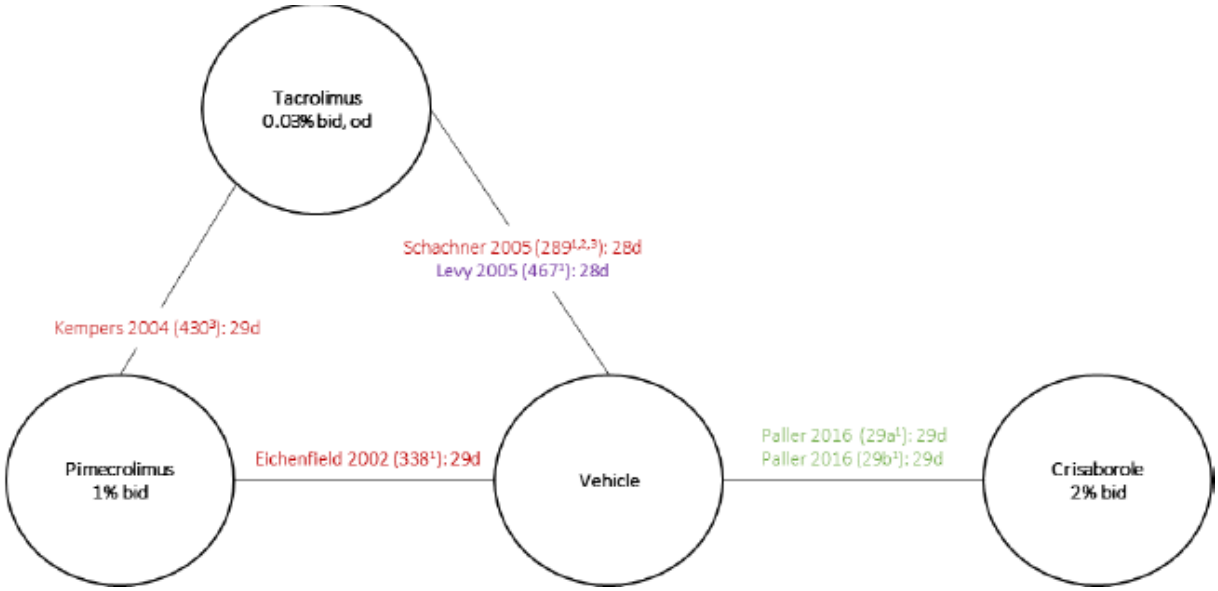
- HRQoL: measured with Short Form (36) Health Survey or EuroQol 5-Dimensions questionnaire.
- Safety: overall adverse events (AEs), discontinuation due to AEs.

Quality Assessment of Included Studies

All included RCTs were evaluated for risk of bias using the Cochrane Risk of Bias tool for RCTs, which summarizes how well each study meets the following five quality criteria: study randomization, concealment of treatment allocation, missing outcome data, blinding of outcome measurement, and completeness of reporting of outcomes, with an overall quality score awarded to each study. There was no description on how the results of risk of bias assessment could have an impact on data analysis.

Evidence Network

Figure 2: Network of Analysis of ISGA Score 0 to 1 at Day 28 to 29



Source: Manufacturer-provided network meta-analysis.[37]

Indirect Comparison Methods

All analyses were conducted within a Bayesian framework using OpenBUGS. All analyses involved a 50,000 run-in iteration phase and a 50,000 iteration phase for parameter estimation. All calculations were performed using OpenBUGS 3.2.2.

The primary end point of interest of this NMA was ISGA score of 0 to 1 at 28 to 29 days. For the ISGA response outcome as well as AEs, there were notable differences across the studies in the size of the baseline risk such as the likelihood of response on vehicle. Given the need to consider baseline risk, a class-effects model was implemented. The authors indicated that such a model would allow for a more stable estimation of the beta (the effect [slope] for baseline risk on relative effects). Specifically, all non-vehicle treatments were grouped into one class, and vehicle into its own class of treatment, with an assumption that non-vehicle treatments varied randomly in efficacy, around a common mean. Fixed or random-effects models (or both, where appropriate) were used for the analyses: class-effects models with baseline risk were fixed-effects for treatment and random-effects for treatments within class, in which all non-vehicle treatments were considered to share a class.

In the analyses, model fit was explored by comparing the deviance information criterion and the posterior mean of the residual deviance for the fixed-effect and random-effect models. Convergence (which is required in Bayesian models in order for the estimates to be valid) was confirmed by evaluating the three-chain Brooks-Gelman-Rubin plots and inspection of the ratios of Monte Carlo error to the standard deviations (SDs) of the posteriors. If convergence was not achieved, then the run-in was increased and/or other factors were examined (such as choice of prior and starting values).

Four analytic models were developed in this study:

- Analysis A: Unadjusted for baseline risk; both random-effect and fixed-effect models were employed; random-effect model used a prior of U[0,1] for the SD of treatment effects across studies within treatment comparisons.
- Analysis B: Baseline risk adjustment was performed only with default prior; only fixed-effect model was used (due to convergence issues); the random-effect model was not run.
- Analysis C: Baseline risk with default prior + class-effects adjustment, prior of U[0,1] for the SD of treatment effects within class; both random-effect and fixed-effect models were employed; an assumption was that all non-standard of care (SOC) treatments share a common relative effect versus SOC, and vary randomly around that effect.
- Analysis D: Baseline risk with default prior + class-effects adjustment, prior of U[0, 0.5] for the SD of treatment effect within class; both random-effect and fixed-effect models were employed; an assumption was that all non-SOC treatments share a common relative effect versus SOC, and vary randomly around that effect.

None of the HRQoL outcomes had sufficient data identified in the systematic literature review for NMA to be feasible. In addition, due to a lack of data, subgroup analyses by age and disease severity were not feasible.

Results

In total, nine RCTs of patients with AD were identified for the NMA, including the AD-301 and AD-302 trials. Treatment durations of these nine RCTs ranged from 28 days to 42 days, and the follow-up duration ranged from 21 days to 43 days. The sample size of these RCTs varied between 133 and 764. Among them, all the included RCTs examined patients older than two years of age, except for one study which enrolled a patient population of one year to 17 years of age and reported an overall average patient age of 6.7 years. One study enrolled patients older than 16 years (average 39.1). Most of the trials (44%) reported on pediatric populations (age range from two years to 17 years), with an additional 33% of studies reporting on a combination of both adult and pediatric patients. For one of the RCTs, the mean age was not reported. The average age of patients in the included RCTs ranged from 6.4 years to 39.1 years. Six RCTs enrolled mixed mild-to-moderate AD populations, while one trial enrolled exclusively mild patients and two trials enrolled exclusively moderate patients.

The interventions in trials of AD were: crisaborole 2% ointment, TCI (pimecrolimus 1% cream; tacrolimus 0.03% or 0.1% ointment), and vehicle. The vehicles used in the included trials were formulated with different emollient properties. None of the included trials reported on the contents of the vehicle, or the proportion of the ingredients. The contents of the study drugs were identified from clinical study reports or drug labels. The contents of the vehicles were assumed to be identical to the contents of the base used for the interventions (white petrolatum, propylene glycol, monoglycerides, diglycerides, paraffin, triglycerides, mineral oil, etc.) by the authors.

The authors reported that the overall risk of bias of the included RCTs was low. The authors indicated that the main reasons for the poor quality of the trials were inadequate treatment allocation concealment and incompleteness of reporting of outcomes. The AD-301 and AD-302 trials were generally considered at low risk of bias.

[redacted]

Atopic Dermatitis Severity

[redacted]

[redacted]1%, or between crisaborole 2% and tacrolimus 0.03% with respect to achieving an ISGA score of 0 to 1 at 28 to 29 days (Table 20).

[redacted]

Table 20: Efficacy Outcome Measures in the Manufacturer's Network Meta-Analysis

	Analysis B (Fixed-Effect Model)	Analysis C (Random-Effect Model)
ISGA Score of 0 to 1 at 28 to 29 Days, OR (95% CrI)		
Crisaborole 2% vs. pimecrolimus 1%	[redacted]	[redacted]
Crisaborole 2% vs. tacrolimus 0.03%	[redacted]	[redacted]
ISGA Score of 0 to 1 at 7 to 8 Days, OR (95% CrI)		
Crisaborole 2% vs. pimecrolimus 1%	[redacted]	[redacted]
Crisaborole 2% vs. tacrolimus 0.03%	[redacted]	[redacted]
Crisaborole 2% vs. tacrolimus 0.1%	[redacted]	[redacted]
ISGA Score of 0 to 1 at 14 to 15 Days, OR (95% CrI)		
Crisaborole 2% vs. pimecrolimus 1%	[redacted]	[redacted]
Crisaborole 2% vs. tacrolimus 0.03%	[redacted]	[redacted]
ISGA Score of 0 to 1 at 21 to 22 Days, OR (95% CrI)		
Crisaborole 2% vs. pimecrolimus 1%	[redacted]	[redacted]
Crisaborole 2% vs. tacrolimus 0.03%	[redacted]	[redacted]
Crisaborole 2% vs. tacrolimus 0.1%	[redacted]	[redacted]
ISGA Score of 0 to 1 at 28 to 43 Days, OR (95% CrI)		
Crisaborole 2% vs. pimecrolimus 1%	[redacted]	[redacted]
Crisaborole 2% vs. tacrolimus 0.03%	[redacted]	[redacted]
Crisaborole 2% vs. tacrolimus 0.1%	[redacted]	[redacted]

CrI = credible interval; ISGA = Investigator's Static Global Assessment; OR = odds ratio.

Median values of the point estimates of the ISGA score were presented.

Source: Manufacturer-provided NMA.[37]

Safety

The authors indicated that an NMA of safety outcomes was not performed because of considerable heterogeneity related to:

1) Inconsistent reporting of data for comparators among included trials and variation in outcome definitions across included trials;
2) It was unclear if an outcome was not reported if it was due to the threshold or definitions or the outcome simply did not occur;
3) Differences in study period between trials (changes in reporting of outcomes data over time; older versus newer trials); and
4) The sparsity of data among included trials.
   Therefore, safety results were qualitatively described in the manufacturer-provided NMA.

The most common AE associated with treatment with crisaborole 2% ointment was application site pain, such as burning or stinging. A less common (less than 1%) AE in patients treated with crisaborole 2% ointment included contact urticaria. The use of TCI was related to local symptoms, such as skin burning (burning sensation, stinging, soreness) or pruritus.

Health-Related Quality of Life

An NMA of HRQoL outcomes was not performed due to the insufficient data identified in the systematic review.

Critical Appraisal

In the manufacturer-provided NMA, the analyses were based on a systematic review of the literature to identify all relevant studies. The literature search included only English-language articles. The methods for study selection and data extraction were suitable. Risk of bias of all individual studies was assessed using the Cochrane Risk of Bias tool, and was generally low as reported by the authors. However, inadequate treatment allocation concealment and incompleteness of reporting of outcomes, as indicated by the authors as the main reasons, could be deemed as a high risk of bias, leading to poor quality of the trial; there was no description of how the results of risk of bias assessment could have an impact on data analysis (e.g., excluding the poor study or conducting a sensitivity analysis, etc.).

Potential sources of heterogeneity with respect to the baseline characteristics were identified, such as age (which ranged from six years to 39 years) and disease severity (six trials enrolled mild-to-moderate patients, one enrolled exclusively mild patients, and two enrolled exclusively moderate patients). However, subgroup analyses by age or baseline disease severity were not conducted, due to the insufficient number of trials within the network. Therefore, there is also uncertainty as to whether relative treatment effects differ by patient age.

In the main report, the primary efficacy outcome in the two pivotal studies was “success in ISGA,” which was defined as ISGA of clear (0) or almost clear (1) with at least a 2-grade improvement from baseline at day 29. In the NMA, the primary efficacy outcome was achieving ISGA score of 0 to 1 at 28 to 29 days. This was a secondary end point from the two pivotal trials. Given that the secondary end point may not be powered to show statistical significance of the difference between treatment groups (type II error), this would make the findings from the NMA more difficult to interpret. If there is statistically significant difference in favour of crisaborole on this secondary end point, we cannot be sure whether observed effect is just a random effect, or whether it is simply due to heterogeneity between trials. [redacted]

TCS is an important treatment option for patients with AD. There is no indirect comparative evidence between crisaborole and TCS such as betamethasone, therefore we are not able to examine the relative efficacy and safety of crisaborole versus TCS in the study population. In addition, there is no indirect comparative evidence for HRQoL or safety outcomes, due to the limited number of trials or sparsity of data.

Treatment durations of the included trials were short (ranging from 28 days to 42 days). The long-term efficacy and safety of crisaborole relative to other topical pharmacological therapies are unknown.

Review of the Incremental Cost-Effectiveness Ratio Review[38]

Objectives and Rationale for the Incremental Cost-Effectiveness Ratio Network Meta-Analysis

To evaluate the comparative clinical effectiveness of crisaborole versus the emollient for management of mild-to-moderate AD.

Methods for the Incremental Cost-Effectiveness Ratio Network Meta-Analysis

This report was based on a review of the literature which included both electronic and manual search components. Multiple databases were searched from January 1996 to January 2017. One single reviewer screened the literature. Eligibility criteria for this study are presented in Table 19. Overall, clinical trials or high-quality systematic reviews of crisaborole or dupilumab compared with emollient or placebo were included. Clinical benefit and safety of the study drugs were assessed. In addition, findings from previously published systematic reviews to inform comparisons of crisaborole to TCS and TCI (with the exception of pimecrolimus, where an NMA was conducted for the comparison between crisaborole and pimecrolimus), and comparisons of dupilumab to cyclosporine, phototherapy, and failed topical therapies were qualitatively reviewed. The criteria published by the US Preventive Services Task Force (USPSTF) were adopted to assess the quality of RCTs and comparative cohort studies.

The findings for the comparisons of dupilumab to the comparators are not presented in this CDR review.

The authors indicated that this review was conducted in accordance with the PRISMA guidelines. The study was not sponsored by the industry.

Data Extraction

It is unclear whether data extraction was performed by two independent reviewers.

Comparators

Crisaborole was planned to compare with emollient, TCS, and TCI. An NMA was conducted to compare crisaborole with pimecrolimus.

Outcomes

Clinical benefits and harms were reported in the study.

Quality Assessment

The criteria published by the US Preventive Services Task Force (USPSTF) were adopted to assess the quality of RCTs and comparative cohort studies.

Evidence Network

Not available.

Meta-Analysis and Indirect Comparison for the Incremental Cost-Effectiveness Ratio Network Meta-Analysis

Crisaborole was evaluated in Studies AD-301 and AD-302 on a 5-point ISGA score. Two other trials (Eichenfield 2002 and Ho 2003) comparing TCI (pimecrolimus) to placebo, using a 6-point ISGA score as an end point were also identified to provide indirect evidence for the comparison between crisaborole and other active comparators. The severity of disease was similar between trials with regard to baseline ISGA score and per cent body surface area involved. Given the lack of head-to-head data and the similar versions of the ISGA score, an indirect comparisons using Bayesian approach was conducted. This analysis was conducted assuming that the clear and almost clear categories were similar on both 5-point and 6-point scales of the ISGA. Random-effect models were used for the analysis.

Results of the Incremental Cost-Effectiveness Ratio Network Meta-Analysis

Data from four trials (AD-301, AD-302, Eichenfield 2002, and Ho 2003) were included in the NMA. The first three trials were also included in the manufacturer-provided NMA. There was no statistically significant difference in efficacy found between crisaborole and pimecrolimus based on ISGA score. The risk ratio for achieving an ISGA score of 0 to 1 between crisaborole and pimecrolimus was 0.61 (95% credible interval [CrI] 0.10 to 2.28; time point not specified).

An NMA was not performed on safety outcomes for the comparison between crisaborole and other active topical therapies.

Critical Appraisal of the Incremental Cost-Effectiveness Ratio Network Meta-Analysis

A systematic review approach was not employed to identify the potentially relevant clinical trials. One single reviewer screened the abstracts and full reports. It was unclear whether data extraction and quality assessment of the included studies were performed by two reviewers. Some important patient baseline characteristics were not described, such as age, prior experience with topical therapies for AD. Subgroup analysis by important patient baseline characteristics, such as age or disease severity was not reported. Therefore, there is also uncertainty on whether relative treatment effects differ by patient age.

In this ICER study, many efficacy outcomes were qualitatively reviewed. An NMA was performed only for the comparison between crisaborole and pimecrolimus. It is unclear why other common topical therapies were not included in the analysis. Similar to the manufacturer-provided NMA, the efficacy outcome in this NMA was the proportion of patients who achieved an ISGA score of 0 to 1. This is a secondary efficacy outcome in AD-301 and AD-302; therefore may not have sufficient power to show statistical significance of the difference between treatment groups.

Similar to the manufacturer-conducted NMA, the included trials in the ICER study were performed in very different time periods (2016 versus 2002 and 2003) and used different versions of ISGA scales (5-point and 6-point). Given the considerable heterogeneity at baseline patient characteristics and the limited number of trials included in the network, it is challenging to make concrete conclusions regarding the relative efficacy of crisaborole and other active treatments.

Conclusion

There are no head-to-head trials comparing crisaborole 2% ointment to other topical pharmacological therapies for patients with mild-to-moderate AD. In the absence of direct evidence, two NMAs comparing crisaborole to other topical pharmacological therapies were identified and summarized for this review. Only one outcome, the proportion of patients achieving an ISGA score of 0 to 1, was assessed in the NMAs, and data were only available for the comparison between crisaborole and TCI (pimecrolimus and tacrolimus). [redacted] However, both NMAs were limited by the number of trials available to inform the network, the fact that only comparisons versus TCI (pimecrolimus and tacrolimus) were reported, that subgroup analyses based on age were not conducted, that there was reporting of only one efficacy outcome (achieving an ISGA score of 0 to 1) to assess comparative treatment effects, and that there was no quantitative assessment of comparative safety. Due to the limitations in the analyses and uncertainty as to whether relative treatment effects differ by patient age, no definitive conclusions regarding the comparative efficacy and safety of crisaborole to other topical therapies can be made for either pediatric or adult patients with mild-to-moderate AD.

References

1. Canadian Dermatology Association. Eczema. 2018; https://dermatology.ca/public-patients/skin/eczema/. Accessed August 28, 2018.
2. Eichenfield LF, Tom WL, Chamlin SL, et al. Guidelines of care for the management of atopic dermatitis: section 1. Diagnosis and assessment of atopic dermatitis. *J Am Acad Dermatol.* 2014;70(2):338-351.
3. Lynde C, Barber K, Claveau J, et al. Canadian practical guide for the treatment and management of atopic dermatitis. *J Cutan Med Surg.* 2005;8 Suppl 5:1-9.
4. Watson W, Kapur S. Atopic dermatitis. *Allergy Asthma Clin Immunol.* 2011;7 Suppl 1:S4.
5. Clinical Study Report: AD-301. A multicenter, randomized, double-blind vehicle -controlled study of the safety and effecacy of AN2728 topical ointment, 2% in children, adolescents, and adults (ages 2 years and older) with atopic dermatitis [**CONFIDENTIAL** internal manufacturer's report]. Palo Alto (CA): Anacor Pharmaceuticals, Inc.; 2015 Nov 23.
6. Clinical Study Report: AD-302. A multicenter, randomized, double-blind vehicle -controlled study of the safety and effecacy of AN2728 topical ointment, 2% in children, adolescents, and adults (ages 2 years and older) with atopic dermatitis [**CONFIDENTIAL** internal manufacturer's report]. Palo Alto (CA): Anacor Pharmaceuticals, Inc.; 2015 Nov 23.
7. Bobotsis R, Fleming P, Eshtiaghi P, Cresswell-Melville A, Drucker AM. A Canadian adult cross-sectional survey of the burden of moderate to severe atopic dermatitis. *J Cutan Med Surg.* 2018;22(4):445-446.
8. Bridgman AC, Eshtiaghi P, Cresswell-Melville A, Ramien M, Drucker AM. The burden of moderate to severe atopic dermatitis in Canadian children: a cross-sectional survey. *J Cutan Med Surg.* 2018;22(4):443-444.
9. 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):116-132.
10. Hoare C, Li Wan Po A, Williams H. Systematic review of treatments for atopic eczema. *Health Technol Assess.* 2000;4(37):1-191.
11. Sidbury R, Davis DM, Cohen DE, et al. Guidelines of care for the management of atopic dermatitis: section 3. Management and treatment with phototherapy and systemic agents. *J Am Acad Dermatol.* 2014;71(2):327-349.
12. American Academy of Dermatology. Atopic dermatitis: Recommendations for the use of systemic immunomodulatory agents. https://www.aad.org/practicecenter/quality/clinical-guidelines/atopic-dermatitis/phototherapy-and-systemic-agents/recommendations-for-systemic-immunomodulatory-agents. Accessed August 28, 2018.
13. Akdis CA, Akdis M, Bieber T, et al. Diagnosis and treatment of atopic dermatitis in children and adults: European Academy of Allergology and Clinical Immunology/American Academy of Allergy, Asthma and Immunology/PRACTALL Consensus Report. *J Allergy Clin Immunol.* 2006;118(1):152-169.
14. e-CPS. *Atopic dermatitis*. Ottawa (ON): Canadian Pharmaceutical Association; 2018: https://www.e-therapeutics.ca. Accessed August 28, 2018.
15. 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.[Erratum appears in J Am Acad Dermatol. 2017 Apr;76(4):777; PMID: 28169014]. *J Am Acad Dermatol.* 2016;75(3):494-503.e496.
16. Paller AS. Clarification of methodology and further results from the pivotal phase 3 study of crisaborole for mild-moderate atopic dermatitis. *Br J Dermatol.* 2018;178(3):663-664.
17. Center for Drug Evaluation and Research. Medical review: Eucrisa (Crisaborole). Anacor Pharmaceuticals, Inc. Application no.: 207695. Approval date: 14/12/2016. Rockville (MD): U.S. Food and Drug Administration (FDA); 2016: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/207695Orig1s000TOC.cfm. Accessed August 28, 2018.
18. Center for Drug Evaluation and Research. Statistical review: Eucrisa (Crisaborole). Anacor Pharmaceuticals, Inc. Application no.: 207695. Approval date: 14/12/2016. Rockville (MD): U.S. Food and Drug Administration (FDA); 2016: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/207695Orig1s000TOC.cfm. Accessed August 28, 2018.
19. CDR submission: Eucrisa (Crisaborole) 2% ointment for topical use. Company: Pfizer Canada, Inc [**CONFIDENTIAL** manufacturer's submission]. Pointe-Claire/Dorval (QU): Pfizer Canada, Inc; 2018 May 25.
20. Hanifin JM, Rajka G. Diagnostic features of atopic dermatitis *Acta Derm Venereol Suppl (Stockh).* 1980;92:44-47.
21. Futamura M, Leshem YA, Thomas KS, Nankervis H, Williams HC, Simpson EL. A systematic review of Investigator Global Assessment (IGA) in atopic dermatitis (AD) trials: many options, no standards. *J Am Acad Dermatol.* 2016;74(2):288-294.
22. Finlay AY, Khan GK. Dermatology Quality of Life Index (DLQI) - a simple practical measure for routine clinical use. *Clin Exp Dermatol.* 1994;19(3):210-216.
23. Basra MK, Fenech R, Gatt RM, Salek MS, Finlay AY. The Dermatology Life Quality Index 1994-2007: a comprehensive review of validation data and clinical results. *Br J Dermatol.* 2008;159(5):997-1035.
24. Lewis-Jones MS, Finlay AY. The Children's Dermatology Life Quality Index (CDLQI): initial validation and practical use. *Br J Dermatol.* 1995;132(6):942-949.

25. Dodington SR, Basra MK, Finlay AY, Salek MS. The Dermatitis Family Impact questionnaire: a review of its measurement properties and clinical application. *Br J Dermatol.* 2013;169(1):31-46.

26. Pfizer Canada Inc response to July 27, 2018 CDR request for additional information regarding the Crisaborole CDR review: Subgroup ananlyses based on baseline disease severity [**CONFIDENTIAL** additional manufacturer's information]. Pointe Claire (QC): Pfizer Canada Inc.; 2018

27. American Academy of Dermatology. Atopic dermatitis clinical guideline, 2014. 2014; https://www.aad.org/practicecenter/quality/clinical-guidelines/atopic-dermatitis. Accessed August 28, 2018.

28. Ahluwalia J, Udkoff J, Waldman A, Borok J, Eichenfield LF. Phosphodiesterase 4 inhibitor therapies for atopic dermatitis: progress and outlook. *Drugs.* 2017;77(13):1389-1397.

29. Clinical Study Report: AD-303. A multicenter, open-label study of the long-term safety of AN2728 topical ointment, 2% in children, adolescents, and adults (ages 2 years and older) with atopic dermatitis [**CONFIDENTIAL** internal manufacturer's report]. Palo Alto (CA): Anacor Pharmaceuticals, Inc.; 2015 Nov 24.

30. Food and Drug Administration. Draft guidance on Pimecrolimus. Silver Spring (MD): FDA; 2012 May.

31. Basra MK, Salek MS, Camilleri L, Sturkey R, Finlay AY. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. *Dermatology.* 2015;230(1):27-33.

32. Yosipovitch G, Simpson EL, Bushmakin AG, et al. Assessment of pruritus in atopic dermatitis. *Itch.* 2018;3(2):e13.

33. Heinl D, Prinsen CAC, Sach T, et al. Measurement properties of quality-of-life measurement instruments for infants, children and adolescents with eczema: a systematic review. *Br J Dermatol.* 2017;176(4):878-889.

34. Lawson V, Lewis-Jones MS, Finlay AY, Reid P, Owens RG. The family impact of childhood atopic dermatitis: the Dermatitis Impact questionnaire. *Br J Dermatol.* 1998;138:107-113.

35. Lewis-Jones MS, Finlay AY, Dykes PJ. The Infants' Dermatitis Quality of Life Index. *Br J Dermatol.* 2001;144:104-110.

36. Eichenfield LF, Call RS, Forsha DW, et al. Long-term safety of crisaborole ointment 2% in children and adults with mild to moderate atopic dermatitis. *J Am Acad Dermatol.* 2017;77(4):641-649.e645.

37. Crisaborole 2% ointment for the treatment of mild-to-moderate atopic dermatitis systematic literuature review and network meta-analysis: final report. *CDR submission: Eucrisa (Crisaborole) 2% ointment for topical use. Company: Pfizer Canada, Inc [**CONFIDENTIAL** manufacturer's submission]*. Dorval, Pointe Claire (QU): Pfizer Canada, Inc.; 2018 May 25.

38. Institute for Clinical and Economic Review. Dupilumab and crisaborole for atopic dermatitis: effectiveness and value: evidence report. Boston (MA): ICER; 2017: https://icer-review.org/wp-content/uploads/2016/10/MWCEPAC_ATOPIC_EVIDENCE_REPORT_051217.pdf. Accessed August 28, 2018.