Common Drug Review
Clinical Review Report

July 2015

Drug	apremilast (Otezla)
Indication	Indicated for the treatment of adult patients with moderate to severe plaque psoriasis who are candidates for phototherapy or systemic therapy.
Listing request	For the treatment of adult patients with moderate to severe plaque psoriasis who have had an inadequate response to, or are intolerant or contraindicated to a conventional systemic therapy.
Manufacturer	Celgene Inc.

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

TABLE OF CONTENTS

Tables

Figures

ABBREVIATIONS

AE	adverse event
BSA	body surface area
CI	confidence interval
DB	double-blind
DLQI	Dermatology Life Quality Index
EQ-5D	EuroQol 5-Dimensions Health-Related Quality of Life Questionnaire
FAS	full analysis set
IL	interleukins
LOCF	last observation carried forward
MCID	minimal clinically important difference
NAPSI	Nail Psoriasis Severity Index
NMA	network meta-analysis
PASI	Psoriasis Area and Severity Index
RCT	randomized controlled trial
SAE	serious adverse event
SD	standard deviation
sPGA	static Physician Global Assessment
TNF	tumour necrosis factor
WDAE	withdrawal due to adverse event

EXECUTIVE SUMMARY

Introduction

Psoriasis is a serious, chronic inflammatory skin disorder that in its worst manifestations may have systemic effects and possibly even be fatal, but more commonly leads to significant symptoms, including pruritus, as well as disfiguring rash that also impairs quality of life. The rash is most often characterized by scaly erythematous papules and plaques, and plaque psoriasis is the most common form of psoriasis. There are approximately one million people who suffer from psoriasis in Canada, and 125 million worldwide. Of these, approximately 90% have plaque psoriasis.[1] Based on a paper by Levy et al., the manufacturer estimated there were approximately 212,500 Canadians with moderate to severe chronic plaque psoriasis.[2]

Treatments for psoriasis range from topical therapies such as corticosteroids to phototherapies to systemic therapies, and these groups can all be used in various combinations with each other. Systemic therapies are further subdivided into established small molecule inhibitors such as methotrexate and cyclosporine, and "biologics," which are all either monoclonal antibodies (e.g., infliximab) or fusion proteins (etanercept). Psoriasis is considered to have a major immune component, and the systemic therapies all target the immune response in various ways. Both the small molecule inhibitors and biologics have significant short- and long-term issues with respect to harms.

Apremilast is a phosphodiesterase-4 (PDE-4) inhibitor, approved for the treatment of moderate to severe plaque psoriasis in patients who are candidates for phototherapy or systemic therapy. This is currently the only indication for apremilast. It is administered orally, and the recommended dose is 30 mg twice daily. The objective of the present review is to perform a systematic review of the beneficial and harmful effects of apremilast for the treatment of moderate to severe plaque psoriasis in patients who are candidates for phototherapy or systemic therapy.

Indication under review
Indicated for the treatment of adult patients with moderate to severe plaque psoriasis who are candidates for phototherapy or systemic therapy.
Listing criteria requested by sponsor
For the treatment of adult patients with moderate to severe plaque psoriasis who have had an inadequate response to, or are intolerant or contraindicated to a conventional systemic therapy.

Results and Interpretation

Included Studies

Two pivotal, phase 3, placebo-controlled, double-blind randomized controlled trials were included in this systematic review. ESTEEM-1 and ESTEEM-2 were multinational studies with Canadian sites, and were sponsored by the manufacturer of apremilast. Both studies had 16-week placebo-controlled phases followed by a 16-week maintenance phase where placebo patients were switched to apremilast, and then a 20-week randomized treatment withdrawal phase where responders were re-randomized to apremilast or placebo. The placebo-controlled phase for both studies was the focus of the systematic review based on pre-specified inclusion criteria; results from the maintenance and withdrawal phases were summarized as supportive evidence.

Enrolled patients had moderate to severe plaque psoriasis for at least 12 months prior to randomization. The primary outcome of both studies was proportion of patients achieving Psoriasis Area and Severity Index (PASI) 75.

Key critical appraisal points included the relatively short duration of the double-blind phase (16 weeks). Apremilast employs a novel mechanism in treating psoriasis; therefore, a longer treatment period would be desirable in order to assess efficacy and, more notably, safety. The lack of an active comparator is also a limitation, as the relative efficacy and safety of apremilast versus key comparators is not known.

Efficacy

Proportion of patients achieving PASI 75 at 16 weeks was the primary outcome of both ESTEEM-1 and ESTEEM-2 (Table 1). Statistically significantly higher proportions of apremilast (33% and 29%) than placebo patients (5% and 6%) achieved PASI 75 at 16 weeks in both ESTEEM-1 (difference in proportions of 27.8% [95% confidence interval (CI), 23.1% to 32.5%], $P < 0.0001$) and in ESTEEM-2 (difference in proportions of 23.0% [95% CI, 16.3% to 29.6%], $P < 0.0001$). Achieving a PASI 75 response is likely a clinically significant outcome; however, it is unclear whether the between-group differences for apremilast versus placebo are clinically significant. The expert on this review believed these between-group differences to be clinically significant. A statistically significant higher proportion of apremilast-treated patients achieved PASI 50 responses compared with placebo patients in both studies; no analysis was planned for PASI 90. Subgroup analyses overall showed consistent results with the main analysis for the proportion of patients achieving PASI 75 at week 16. In ESTEEM-1, PASI 75 results for apremilast versus placebo were similar whether patients had a history of psoriatic arthritis ([redacted]% versus [redacted]%, respectively) or not ([redacted]% versus [redacted]%). Similar results for PASI 75 were seen in ESTEEM-2, with no obvious difference in response between patients with a history of psoriatic arthritis ([redacted]% versus [redacted]%) and those without ([redacted]% versus [redacted]%).

Quality of life was assessed using the disease-specific Dermatology Life Quality Index (DLQI), as well as generic instruments such as the EuroQol 5-Dimensions Health-Related Quality of Life Questionnaire (EQ-5D) (Table 1). The EQ-5D results were specified as an exploratory analysis by the manufacturer; therefore, no adjustment for multiplicity was made and these EQ-5D results may be susceptible to inflated type 1 error. Apremilast improved quality of life at 16 weeks on the DLQI compared with placebo in both ESTEEM-1 (difference in least squares [LS] means of –4.5 [95% CI, –5.4 to –3.6], $P < 0.0001$) and ESTEEM-2 (difference in LS means of –4.0 [95% CI, –5.3 to –2.8], $P < 0.0001$). These improvements were both statistically and clinically significant, based on the minimal clinically important difference (MCID) for the DLQI of 3.2. [redacted]

Among other key efficacy outcomes (Table 1), a larger proportion of apremilast patients achieved a static Physician Global Assessment (sPGA) response after 16 weeks compared with placebo, and this difference was statistically significant in ESTEEM-1 (difference in proportions of 17.8% [95% CI, 13.7 to 21.9], $P < 0.0001$) and in ESTEEM-2 (difference in proportions of 16.1% [95% CI, 10.2 to 21.9],

P < 0.0001). The affected body surface area (BSA) was also reduced to a greater extent by apremilast compared with placebo at 16 weeks, and this difference was statistically significant in both ESTEEM-1 (difference in LS means of –40.78% [–46.34 to –35.21], *P* < 0.0001) and in ESTEEM-2 (difference in LS means of –42.15% [–51.11 to –33.20], *P* < 0.0001). The MCIDs for sPGA and affected BSA are not known; however, the clinical expert consulted for the review believed these to be clinically significant differences.

Other efficacy outcomes included symptom measures, including the pruritus visual analogue scale (VAS). In ESTEEM-1 (difference in LS means [two-sided 95% CI] of –24.2 [–28.7 to –19.8], *P* < 0.0001) and ESTEEM-2 (–21.3 [–28.4 to –14.2], *P* < 0.0001) there was a statistically significant improvement in pruritus after 16 weeks for apremilast versus placebo. Nail symptoms were also assessed using the Nail Psoriasis Severity Index (NAPSI). NAPSI scores improved from baseline to week 16 for apremilast versus placebo in both ESTEEM-1 (difference in LS means [two-sided 95% CI] –29.1 [–28.6 to –16.5], *P* < 0.0001) and ESTEEM-2 (difference in LS means [two-sided 95% CI] –22.9 [–38.9 to –6.9], *P* = 0.0052).

Data from the maintenance phase (weeks 16 through 32), in which all patients received apremilast, suggest that the effects of apremilast persisted with the additional 16 weeks of therapy in the apremilast group (28% of patients with PASI 75 in ESTEEM-1, 25% in ESTEEM-2), while the group that switched from placebo to apremilast had similar responses to those receiving 32 weeks of apremilast (31% and 29%, respectively) (Table 11). These findings are limited, however, by the fact that at this point in the study, all patients would know that they are on apremilast; therefore, blinding is lost. In the re-randomization phase, more patients who were switched to placebo lost their PASI 75 responses compared with those who continued on apremilast; however, this is not a surprising finding, given active treatment withdrawal design of this phase of the study (Table 12).

The manufacturer submitted a network meta-analysis (NMA) based on data from a systematic review it performed that included all key comparators, such as cyclosporine and methotrexate. However, the statistical analyses that were performed were versus placebo, rather than comparing apremilast to other active comparators, and this limits the conclusions that can be drawn from this analysis. The NMA focused solely on PASI (PASI 50/75/90), and did not report on harms. [redacted]

Harms

In ESTEEM-1, 69% of apremilast patients and 56% of placebo patients reported an adverse event (AE) after 16 weeks of therapy, while in ESTEEM-2, 68% of apremilast patients and 60% of placebo patients experienced an AE (Table 1). The most common AEs were diarrhea (18% of apremilast patients versus 7% of placebo patients across studies) and nausea (17% of apremilast patients and 7% placebo patients). Apremilast is associated with gastrointestinal adverse effects. Serious adverse events (SAEs) were reported in 2% of apremilast patients and 3% of placebo patients after 16 weeks in ESTEEM-1, and in 2% of each of the apremilast and placebo groups in ESTEEM-2. No single SAE occurred in more than a single patient. Withdrawals due to adverse event (WDAEs) occurred in 5% of apremilast patients and 3% of placebo patients after 16 weeks in ESTEEM-1, and 6% and 5% of patients, respectively, in ESTEEM-2. The most common reason for withdrawal across groups was nausea. Aside from gastrointestinal adverse effects, weight loss was a notable harm associated with apremilast, and the proportion of patients with weight loss as an AE was less than 2% in the apremilast groups and less than 1% in the placebo groups of both ESTEEM studies. Of note, however, the proportion of patients with clinically significant weight loss (i.e., weight decreases of > 5% of baseline body weight), not specified as an AE, was [redacted]% for apremilast

versus [redacted]% for placebo in ESTEEM-1, and [redacted]% versus [redacted]%, respectively, in ESTEEM-2. No additional safety issues were detected in the maintenance (Table 13) or re-randomization phases (Table 14), with follow-up to 52 weeks, or the extension, with follow-up to 104 weeks (Table 16).

Conclusions

Two multinational, manufacturer-sponsored, pivotal, phase 3, double-blind randomized controlled trials, ESTEEM-1 and ESTEEM-2, met the inclusion criteria for this review. Both studies, ESTEEM-1 (N = 844) and ESTEEM-2 (N = 413) randomized patients with moderate to severe chronic plaque psoriasis in a 2:1 manner to either apremilast 30 mg twice daily or placebo, over a 16-week double-blind comparison period. Apremilast was statistically significantly superior to placebo after 16 weeks for proportion of patients achieving a PASI 75, the primary outcome of both studies. Apremilast also improved the sPGA responses versus placebo, and reduced the affected BSA versus placebo, all statistically significant differences. Quality of life was statistically and clinically significantly improved after 16 weeks versus placebo in both studies using a dermatology-specific instrument, the DLQI. The most common AEs with apremilast were gastrointestinal-related nausea and diarrhea, and these were also the most common reasons for WDAEs. There were no clear indications of any serious harms issues with apremilast, although interpretation of this finding is limited somewhat by the relatively short 16-week follow-up in the double-blind comparative phase. A manufacturer-submitted NMA suggested the biologics provided superior efficacy to the small molecule drugs for psoriasis, including apremilast, based on rank ordering alone. There is uncertainty with respect to the conclusions from the NMA because the manufacturer did not provide statistical indirect comparison estimates.

TABLE 1: SUMMARY OF RESULTS

	ESTEEM-1		ESTEEM-2	
PASI 75	**Apremilast N = 562**	**Placebo N = 282**	**Apremilast N = 274**	**Placebo N = 137**
Patients at week 16, n (%) — FAS/LOCF	186 (33)	15 (5)	79 (29)	8 (6)
Difference in proportions (95% CI)[a]	27.8 (23.1 to 32.5)		23.0 (16.3 to 29.6)	
P value	*P* < 0.0001		*P* < 0.0001	
NRI[b]/LOCF	183 (33)	14 (5)	79 (29)	7 (5)
sPGA				
sPGA response, week 16, n (%)[c]	122 (22)	11 (4)	56 (20)	6 (4)
Difference in proportions [95% CI][a]	17.8 [13.7 to 21.9]		16.1 (10.2 to 21.9)	
P value	*P* < 0.0001		*P* < 0.0001	
Affected BSA				
Mean (SD) at baseline	24.4 (14.7)	25.0 (14.3)	25.5 (15.5)	27.6 (15.9)
Mean (SD) % change at week 16	–47.8 (38.5)	–6.9 (38.9)	–48.5 (40.8)	–6.1 (47.6)
LS mean (95% CI)	–47.8 (–51.0 to –44.6)	–7.0 (–11.5 to –2.4)	–48.4 (53.6 to –43.2)	–6.3 (–13.5 to 1.1)
Difference in LS means (2-sided 95% CI)[d]	–40.78 (–46.34 to –35.21)		–42.15 (–51.11 to –33.20)	
P value	*P* < 0.0001		*P* < 0.0001	
DLQI Total Score				
Mean (SD) at baseline	12.7 (7.1)	12.1 (6.7)	12.5 (7.1)	12.8 (7.1)
Mean (SD) change at week 16	–6.6 (–6.7)	–2.1 (5.7)	–6.7 (7.0)	–2.8 (7.2)
LS mean change [95% CI]	–6.6 (–7.1 to –6.1)	–2.1 (–2.8 to –1.3)	–6.7 (–7.5 to –6.0)	–2.7 (–3.7 to –1.7)
Difference in LS means (2-sided 95% CI)	–4.5 (–5.4 to –3.6)[e]		–4.0 (–5.3 to –2.8)[f]	
P value	*P* < 0.0001		*P* < 0.0001	
EQ-5D VAS				
Mean (SD) at baseline	[redacted] N = [redacted]	[redacted] N = [redacted]	[redacted] N = [redacted]	[redacted] N = [redacted]
Mean (SD) change at week 16	[redacted]	[redacted]	[redacted]	[redacted]
LS mean change from baseline (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Difference in LS means (2-sided 95% CI)[f]	[redacted]		[redacted]	
2-sided *P* value	*P* = [redacted]		*P* = [redacted]	
EQ-5D Index Value Score				
Mean (SD) at baseline	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD) change at week 16	[redacted]	[redacted]	[redacted]	[redacted]
LS mean change from baseline (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Difference in LS means (2-sided 95% CI)[g]	[redacted]		[redacted]	
2-sided *P* value	*P* [redacted]		*P* [redacted][h]	

	ESTEEM-1		ESTEEM-2	
PASI 75	**Apremilast N = 562**	**Placebo N = 282**	**Apremilast N = 274**	**Placebo N = 137**
Serious Adverse Events				
Patients, n (%)	12 (2)	8 (3)	5 (2)	3 (2)
Adverse Events				
Patients, n (%)	388 (69)	157 (56)	185 (68)	82 (60)
WDAE				
WDAEs, n (%)	29 (5)	9 (3)	15 (6)	7 (5)

BSA = body surface area; CI = confidence interval; DLQI = Dermatology Life Quality Index; EQ-5D = EuroQol 5-Dimensions Health-Related Quality of Life Questionnaire; FAS = full analysis set; LOCF = last observation carried forward; LS = least squares; NRI = non-responder imputation; PASI = Psoriasis Area and Severity Index; SD = standard deviation; sPGA = static Physician Global Assessment; VAS = visual analogue scale; WDAE = withdrawal due to adverse event.

[a] Two-sided 95% CI is based on the normal approximation. Two-sided *P* value is based on the two-sided chi-square test.

[b] This is a sensitivity analysis using both LOCF and NRI to account for missing values. NRI patients who discontinued before week 16 or whose PASI evaluations were missing were counted as non-responders.

[c] sPGA response: score of 0 or 1 with ≥ two-point reduction from baseline.

[d] Based on an analysis of covariance model for the per cent change from baseline at week 16, with treatment group as a factor and the baseline value as a covariate. Means (LS means), difference in LS means, and *P* values were adjusted by covariate. The two-sided *P* value for slope homogeneity is > 0.05.

[e] Based on an analysis of variance model for the change from baseline at week 16, with treatment group as a factor (an analysis of variance [ANOVA] model). Unadjusted means and *P* values are provided. The two-sided *P* value for slope homogeneity is < 0.05.

[f] Based on an analysis of covariance model for the per cent change from baseline at week 16, with treatment group as a factor and the baseline value as a covariate. Means (LS means) and *P* values were adjusted by covariate. The two-sided *P* value for slope homogeneity is > 0.05.

[g] Based on an analysis of covariance model for the change from baseline at week 16. The model includes treatment group as a factor (ANOVA model). The unadjusted means and *P* value are provided. The two-sided *P* value for slope homogeneity is < 0.05.

[h] *P* value was ≤ 0.05 and considered nominally significant, as there was no adjustment for multiplicity based on hierarchical testing.

Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

1. INTRODUCTION

1.1 Disease Prevalence and Incidence

Psoriasis is a serious, chronic inflammatory skin disorder that in its worst manifestations may have systemic effects and possibly even be fatal, but more commonly leads to significant symptoms, including pruritus, as well as having impacts on appearance and reduced quality of life. The appearance is most often characterized by scaly erythematous papules and plaques, with plaque psoriasis being the most common form of psoriasis. Plaque psoriasis is characterized by well-demarcated papules that are covered by silvery scales. Moderate to severe psoriasis is defined by extent of skin coverage (involvement of more than 5% to 10% of body surface area [BSA]), or location (involvement of the face, palm, or sole), or severity (disease that is disabling). Once patients have exceeded this 5% to 10% of their skin involvement, topical therapy becomes more problematic for the patient, as there is such a large surface area to cover. At this point these patients tend to move to systemic therapy (see section 1.2, Standards of Therapy).[5,6]

There are approximately one million people who suffer from psoriasis in Canada, and 125 million worldwide. Of these, approximately 90% have plaque psoriasis.[1] Based on a paper by Levy et al., the manufacturer estimated there were approximately 212,500 Canadians with moderate to severe chronic plaque psoriasis.[2]

1.2 Standards of Therapy

Psoriasis is treated topically, including phototherapy, and with systemic therapies, often administered concomitantly. Topical therapies are often corticosteroids of varying potencies; however, emollients, coal tar, vitamin D analogues, and topical retinoids may also be used. Phototherapies may either be strictly topical (ultraviolet blue [UVB] light on involved skin) or combine a systemic drug like psoralen with phototherapy. The fact that topical therapies are applied locally is both an advantage with respect to reduced risk of harms, but also a disadvantage in that widely disseminated lesions will require large amounts of topical therapy, creating an added burden for the patient. Psoriasis is essentially an immune disorder; therefore, the systemic therapies all work by suppressing components of the immune system. The first systemic therapies, often referred to as “conventional therapies,” were all small molecules, the two most important currently being methotrexate, an antimetabolite also used in some cancers and rheumatoid arthritis, and cyclosporine, a potent immunosuppressant also used in prevention of organ transplant rejection. Both of these drugs have significant toxicities associated with them. The biologics, monoclonal antibodies, and fusion proteins were the next systemic therapies to be developed, and all of these original biologics targeted tumour necrosis factor (TNF), a key mediator of inflammation. The newest biologics, both monoclonal antibodies, target interleukins (IL). Ustekinumab blocks IL-12 and IL-23, and secukinumab, currently under review by Health Canada, blocks IL-17. The TNF inhibitors have all been associated with elevated risk of certain cancers with long-term use and increased risk of infection, including tuberculosis. The association between TNF inhibitor use and increased risk of cancer is less well defined and more controversial in psoriasis, according to the clinical expert. Ustekinumab has also been associated with increased infection risk and malignancy, and more recently, serious skin reactions.[6]

1.3 Drug

Apremilast is an orally administered phosphodiesterase-4 (PDE-4) inhibitor, given at a dose of 30 mg twice daily. Apremilast is under review for the treatment of adult patients with moderate to severe plaque psoriasis who have had an inadequate response or are intolerant or contraindicated to a conventional systemic therapy. This is currently the only approved indication for apremilast.

Indication under review
Indicated for the treatment of adult patients with moderate to severe plaque psoriasis who are candidates for phototherapy or systemic therapy.
Listing criteria requested by sponsor
For the treatment of adult patients with moderate to severe plaque psoriasis who have had an inadequate response to, or are intolerant or contraindicated to a conventional systemic therapy.

TABLE 2: KEY CHARACTERISTICS OF SMALL MOLECULE INHIBITORS AND BIOLOGICS

	Apremilast	Cyclosporine	Methotrexate	
Mechanism of Action	PDE-4 inhibitor	Calcineurin inhibitor, inhibits IL-2, preventing T-cell activation	Antimetabolite; folate antagonist	
Indication[a]	Moderate to severe plaque psoriasis in patients who are candidates for phototherapy or systemic therapy	Psoriasis	Psoriasis	
Route of Administration	Oral	Oral	Oral	
Recommended Dose	30 mg twice daily	• 2.5 mg/kg/day given in 2 divided oral doses, 12 hours apart • Dose may be titrated to achieve effect • Total daily dose should not exceed 5 mg/kg/day	• Weekly single oral, IM or IV dose schedule: 10 mg to 25 mg per week until adequate response is achieved. • Dosages in each schedule may be gradually adjusted to achieve optimal clinical response; 30 mg/week should not ordinarily be exceeded.	
Serious Side Effects/Safety Issues	No serious harms detected in RCTs with 16 weeks follow-up	• Infections • Nephrotoxicity • Hypertension	• Bone marrow suppression • Hepatotoxicity • Nephrotoxicity • Alopecia • Stomatitis	

	Infliximab	Adalimumab	Etanercept	Ustekinumab
Mechanism of Action	• TNF inhibitor • Chimeric monoclonal antibody	• TNF inhibitor • Recombinant human monoclonal antibody	• TNF inhibitor • Fusion protein	• IL-12 and IL-23 inhibitor • Fully human monoclonal antibody
Indication[a]	Chronic moderate to severe plaque psoriasis who are candidates for systemic therapy. For patients with chronic moderate plaque psoriasis, should be used after phototherapy has been shown to be ineffective or inappropriate.	Chronic moderate to severe psoriasis who are candidates for systemic therapy. For patients with chronic moderate plaque psoriasis, adalimumab should be used after phototherapy has been shown to be ineffective or inappropriate.	Chronic moderate to severe plaque psoriasis in patients who are candidates for systemic therapy or phototherapy.	Chronic moderate to severe plaque psoriasis in patients who are candidates for phototherapy or systemic therapy.
Route of Administration	Intravenous	Subcutaneous	Subcutaneous	Subcutaneous
Recommended Dose	5 mg/kg given as an intravenous infusion followed by additional 5 mg/kg doses at 2 and 6 weeks after the first infusion, then every 8 weeks thereafter. If a patient does not show an adequate response at week 14, after infusions at weeks 0, 2, and 6, no additional treatment with infliximab should be given.	80 mg administered subcutaneously, followed by 40 mg subcutaneously given every other week starting 1 week after the initial dose. Continued therapy beyond 16 weeks should be carefully reconsidered in a patient not responding within this time period.	50 mg dose given twice weekly (administered 3 or 4 days apart) for 3 months followed by a reduction to a maintenance dose of 50 mg per week. A maintenance dose of 50 mg given twice weekly has also been shown to be efficacious.	45 mg at weeks 0 and 4, then every 12 weeks thereafter. Alternatively, 90 mg may be used in patients with a body weight greater than 100 kg. For patients who inadequately respond to dosing every 12 weeks, consideration may be given to treating as often as every 8 weeks. Consider discontinuing treatment in patients who have shown no response up to 12 weeks of treatment.
Serious Side Effects/Safety Issues	• Infection • Cancer	• Infection • Cancer	• Infection • Cancer	• Infection • Cancer • Serious skin reactions (exfoliative dermatitis and erythrodermic psoriasis)

IL = interleukin; IM = intramuscular; IV = intravenous; PDE-4 = phosphodiesterase-4; RCTs = randomized controlled trials; TNF = tumour necrosis factor.
[a] Health Canada indication.
Source: e-CPS.[7]

2. OBJECTIVES AND METHODS

2.1 Objectives

To perform a systematic review of the beneficial and harmful effects of apremilast for the treatment of moderate to severe plaque psoriasis in patients who are candidates for phototherapy or systemic therapy.

2.2 Methods

Studies were selected for inclusion in the systematic review based on the pivotal studies submitted by the manufacturer and the selection criteria presented in Table 3.

TABLE 3: INCLUSION CRITERIA FOR THE SYSTEMATIC REVIEW

Patient Population	Adult patients with moderate to severe plaque psoriasis who are candidates for phototherapy or systemic therapy. Subgroups: patients with psoriatic arthritis; patients with inadequate response to, who are intolerant of, or have contraindications to systemic therapy.
Intervention	Apremilast 30 mg twice daily alone or in combination with other drug or non-drug therapies for moderate to severe plaque psoriasis.
Comparators	As monotherapy or in combination: **Systemic:** • Methotrexate • Cyclosporine • Acitretin • Etanercept • Infliximab • Adalimumab • Ustekinumab **Topical:** • Tazarotene • Vitamin D analogues (e.g., calcitriol, calcipotriol) • Topical corticosteroids
Outcomes	**Key efficacy outcomes:** • Quality of life (DLQI, EQ-5D) • Psoriatic Area Severity Index • Physician Global Assessments (e.g., scalp, palmoplantar) • Proportion of body surface area involved **Other efficacy outcomes:** • Other symptoms (e.g., pruritus, nail) **Harms outcomes:** • AEs • SAEs • WDAEs • Notable harms (neuropsychiatric effects, weight loss, gastrointestinal)
Study Design	Published and unpublished phase 3 RCTs

AE = adverse event; DLQI = Dermatology Life Quality Index; EQ-5D = EuroQol 5-Dimensions Health-Related Quality of Life Questionnaire; RCT = randomized controlled trial; SAE = serious adverse event; WDAE = withdrawal due to adverse event.

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 (1946–) with in-process records and daily updates through Ovid; Embase (1974–) through 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 concept was Otezla (apremilast).

No filters were applied to limit the 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.

The initial search was completed on December 8, 2014. Regular alerts were established to update the search until the meeting of the CADTH Canadian Drug Expert Committee (CDEC) on April 8, 2015. 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 (http://www.cadth.ca/en/resources/finding-evidence-is/grey-matters): health technology assessments, health economics, clinical practice guidelines, drug regulatory approvals, advisories and warnings, drug class reviews, clinical trials, 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 CADTH Common Drug Review (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: EXCLUDED STUDIES.

3. RESULTS

3.1 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 2 and described in Section 3.2. A list of excluded studies is presented in APPENDIX 3: EXCLUDED STUDIES.

FIGURE 1: QUOROM FLOW DIAGRAM FOR INCLUSION AND EXCLUSION OF STUDIES

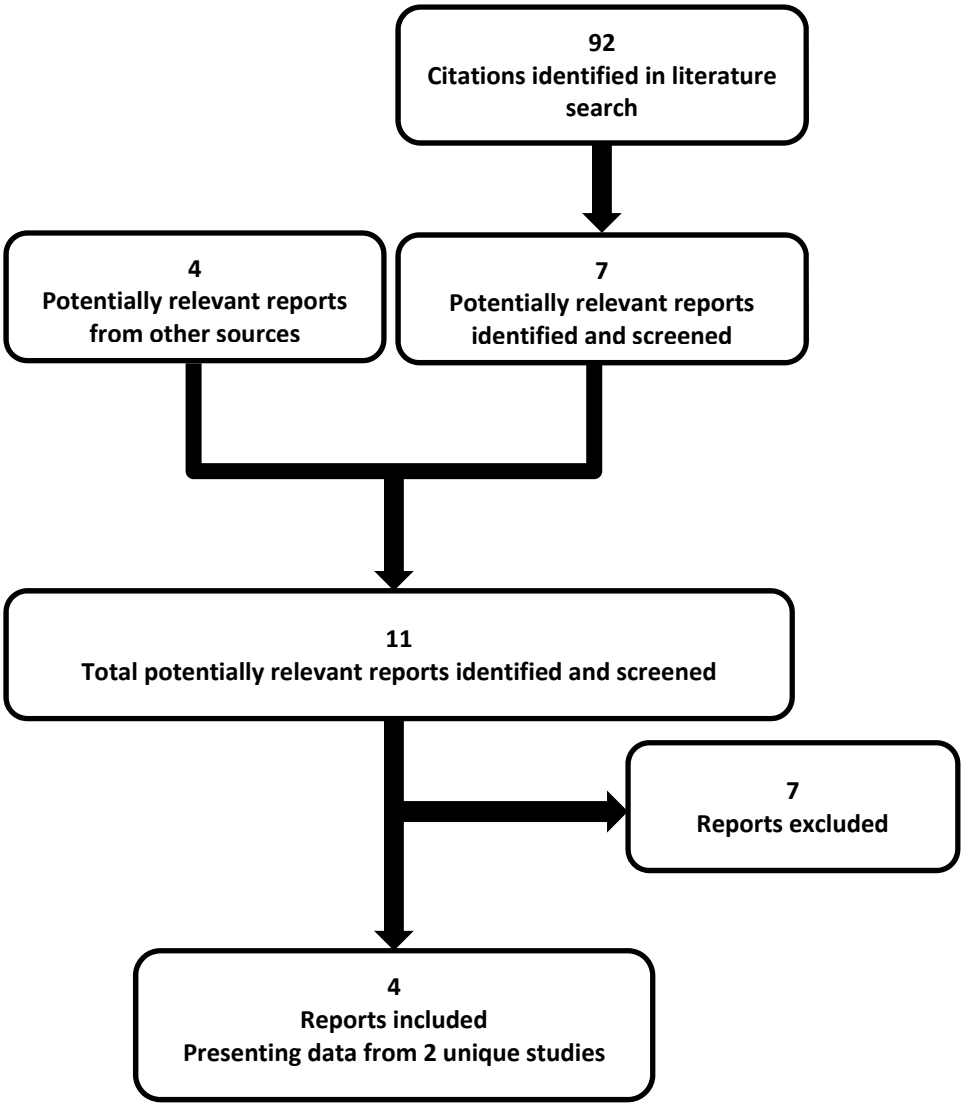


TABLE 4: DETAILS OF INCLUDED STUDIES

		ESTEEM-1	ESTEEM-2
DESIGNS & POPULATIONS	**Study Design**	DB RCT	DB RCT
	Locations	72 sites: Canada, US, EU, Australia	40 sites: Canada, US, EU
	Randomized (N)	N = 844	N = 413

	Inclusion Criteria	• Diagnosis of chronic plaque psoriasis for at least 12 months prior to screening
		• Had moderate to severe plaque psoriasis at screening and baseline as defined by:
		  ○ PASI score ≥ 12
		  ○ BSA ≥ 10%
		  ○ sPGA ≥ 3 (moderate)
		• Was a candidate for phototherapy and/or systemic therapy	

	Exclusion Criteria	• Psoriasis flare or rebound within 4 weeks prior to screening • Topical therapy within 2 weeks of randomization • Patients with scalp psoriasis were permitted to use coal tar shampoo and/or salicylic acid scalp preparations on scalp lesions. An unmedicated skin moisturizer (e.g., Eucerin) was permitted for body lesions only. Patients should not have used these topical treatments within 24 hours prior to the clinic visit • Systemic therapy for psoriasis within 4 weeks prior to randomization • Use of phototherapy within 4 weeks prior to randomization • Adalimumab, etanercept, efalizumab, infliximab, or certolizumab pegol within 12 weeks prior to randomization • Alefacept, briakinumab, or ustekinumab within 24 weeks prior to randomization	
DRUGS	**Intervention**	Apremilast 30 mg twice daily	
	Comparator(s)	Placebo	
DURATION	**Phase**		
	Run-in	Maximum 35 days	
	Double-blind	16 weeks	
	Follow-up	16 week maintenance; 20-week withdrawal phase; extension 208 weeks	
OUTCOMES	**Primary End Point**	Proportion of patients who achieved at least a 75% reduction in PASI (PASI 75) at week 16 from baseline.	
	Other End Points	• Patients with an sPGA score of clear (0) or almost clear (1) with at least 2-point reduction from baseline at week 16 • Change from baseline in the pruritus VAS at week 16 • Change from baseline in the DLQI total score at week 16 • Change from baseline in the MCS score of SF-36 at week 16	• Per cent change from baseline in psoriasis affected BSA (%) at week 16 • Per cent change from baseline in the PASI score at week 16 • Proportion of patients who achieve PASI 50 at week 16 • Change from baseline in the Pruritus VAS at week 16 • Change from baseline in the DLQI

		ESTEEM-1	ESTEEM-2
		• Proportion of patients who achieved both PASI 75 and sPGA score of clear (0) or almost clear (1) with at least 2-point reduction from baseline at week 16 • Time to loss of PASI 75 response (loss of effect) during the randomized treatment withdrawal phase	total score at week 16 • Change from baseline in the MCS score of SF-36 at week 16 • Proportion of patients who achieved both PASI 75 and sPGA score of clear (0) or almost clear (1) with at least 2 points reduction from baseline at week 16 • Time to loss of effect (i.e., loss of 50% of the improvement in PASI score obtained at week 32 compared with baseline) during the randomized treatment withdrawal phase
NOTES	**Publications**	None	None

BSA = body surface area; DB = double-blind; DLQI = Dermatology Life Quality Index; MCS = mental component summary; PASI = Psoriasis Area and Severity Index; RCT = randomized controlled trial; SF-36 = Short-Form 36-Item Health Survey; sPGA = static Physician Global Assessment; VAS = visual analogue scale.
Note: Four additional reports were included (Clinical Study Reports for ESTEEM-1 and ESTEEM-2[3,4]; Health Canada Review Report[8]; manufacturer's submission[9]).
Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

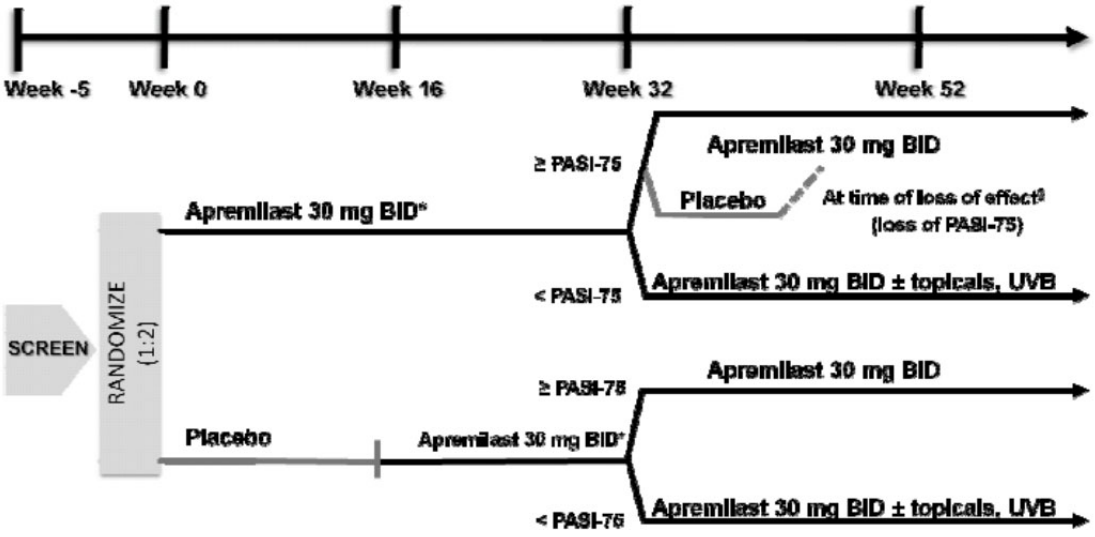
3.2 Included Studies

3.2.1 Description of Studies

ESTEEM-1 and ESTEEM-2 were both placebo-controlled, double-blind (DB) randomized controlled trials (RCTs) with three different phases (Table 4). In each study, patients were randomized 2:1 to either apremilast or placebo, and stratification was not reported. The initial DB comparison to placebo lasted 16 weeks, and was followed by a 16-week maintenance phase where patients originally assigned to apremilast remained on the drug, while patients originally assigned to placebo were switched to apremilast. Finally, weeks 32 to 52 were referred to as the randomized treatment withdrawal phase, and tested the durability of response to apremilast. At week 32, responders (those achieving Psoriasis Area and Severity Index [PASI] of at least 75 in ESTEEM-1 and 50 in ESTEEM-2) were re-randomized to either continue on apremilast or switch to placebo. Patients experiencing a loss of PASI 75 during this phase were to resume apremilast. Patients who did not achieve a PASI 75 were allowed to add topical therapies and/or phototherapy to their apremilast regimen at week 32. This decision to add therapy could be made only at week 32 and not thereafter. The long-term extension phase started at week 52 and continued for 208 weeks, for a total study duration of up to five years. The focus of this review was the initial 16-week DB comparison to placebo.

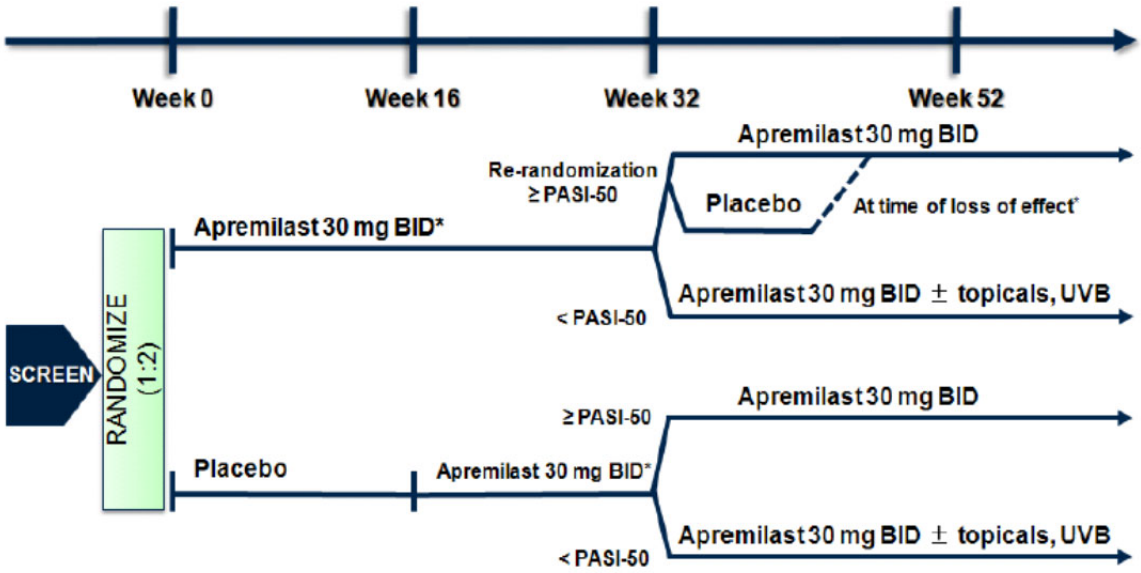
FIGURE 2: SCHEMATIC OF STUDY DESIGN FOR ESTEEM-1 AND ESTEEM-2.

a) ESTEEM-1



BID = twice daily; PASI 75 = Psoriasis Area and Severity Index 75; UVB = ultraviolet B light.

b) ESTEEM-2



BID = twice daily; PASI 50 = Psoriasis Area and Severity Index 50; UVB = ultraviolet B light.
Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

3.2.2 Populations

a) Inclusion and exclusion criteria

Patients were to have moderate to severe plaque psoriasis defined by PASI score (≥ 12), BSA (≥ 10%), and static Physician Global Assessment (sPGA) (≥ 3), and had to have chronic plaque psoriasis for at least 12 months (Table 4). Patients were excluded if they had a psoriasis flare or rebound within four weeks prior to screening. Patients were also excluded if they had systemic or phototherapy within four weeks of randomization, and biologics either 12 weeks (adalimumab, etanercept, efalizumab, infliximab, or certolizumab pegol) or 24 weeks (alefacept, briakinumab, or ustekinumab) prior to randomization.

b) Baseline characteristics

Patients across ESTEEM-1 and ESTEEM-2 were 46 years of age, on average, and the majority were male (approximately 70%) (Table 5). The majority (approximately 90%) were Caucasian across both studies. On average, patients had plaque psoriasis for approximately 19 years. The mean total PASI score ranged from 18.7 to 20.0 points across both studies, and approximately one-third of patients had PASI scores > 20. The mean BSA involvement ranged from 24.4% to 27.6%, with about half of patients above a BSA of 20%.

There were some differences between groups in each study with respect to baseline demographics. There appeared to be a lower proportion of males in the apremilast group in ESTEEM-2 compared with placebo, a lower proportion of apremilast patients who had no prior phototherapy, and had failed prior TNF. In ESTEEM-2, 15% of apremilast patients had psoriatic arthritis at baseline compared with 10% of placebo patients. Baseline PASI scores were similar between groups in each study, while BSA involvement was greater than 20% in fewer apremilast patients than placebo patients: 52% versus 58%.

TABLE 5: SUMMARY OF BASELINE CHARACTERISTICS

	ESTEEM-1		ESTEEM-2	
	Apremilast N = 562	**Placebo N = 282**	**Apremilast N = 274**	**Placebo N = 137**
Mean (SD) Age, Years	45.8 (13.1)	46.5 (12.7)	45.3 (13.1)	45.7 (13.4)
Male Gender, n (%)	379 (67)	194 (69)	176 (64)	100 (73)
Race				
White	[redacted]	[redacted]	[redacted]	[redacted]
Asian	[redacted]	[redacted]	[redacted]	[redacted]
Black/African American	[redacted]	[redacted]	[redacted]	[redacted]
Native American/Pacific Islander	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]
Weight, mean (SD) kg	93.2 (21.4)	93.7 (23.2)	91.4 (23.0)	90.5 (22.5)
Duration of Plaque Psoriasis, Mean (SD) Years	19.8 (13.0)	18.7 (12.4)	17.9 (11.4)	18.7 (12.1)
History of Guttate-erythrodermic Psoriasis or Sudden Intensification in Psoriasis	[redacted]	[redacted]	[redacted]	[redacted]
Scalp psoriasis	[redacted]	[redacted]	[redacted]	[redacted]
Nail psoriasis	[redacted]	[redacted]	[redacted]	[redacted]
Palmoplantar psoriasis	[redacted]	[redacted]	[redacted]	[redacted]
Psoriatic arthritis	[redacted]	[redacted]	[redacted]	[redacted]
Total PASI Score				
Mean (SD)	18.7 (7.2)	19.4 (7.4)	18.9 (7.1)	20.0 (8.0)
≤ 20%	[redacted]	[redacted]	[redacted]	[redacted]
> 20%	158 (28)	87 (31)	81 (30)	49 (36)
BSA Involvement				
Mean (SD) %	24.4 (14.7)	25.3 (14.6)	25.5 (15.4)	27.6 (15.8)
≤ 20	296 (53)	133 (47)	131 (48)	57 (42)
> 20	266 (47)	149 (53)	143 (52)	80 (58)
Number of Prior Phototherapies				
0	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]
≥ 3	[redacted]	[redacted]	[redacted]	[redacted]
Number of Failed Prior Phototherapies				
0	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]
≥ 3	[redacted]	[redacted]	[redacted]	[redacted]
Number of Prior Conventional Systemic Therapies				
0	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]
≥ 3	[redacted]	[redacted]	[redacted]	[redacted]

	ESTEEM-1		ESTEEM-2	
Number of Failed Prior Conventional Systemic Therapies				
0	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]
≥ 3	[redacted]	[redacted]	[redacted]	[redacted]
Number of Prior Biologic Therapies				
0	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]
≥ 3	[redacted]	[redacted]	[redacted]	[redacted]
Number of Failed Prior Biologic Therapies				
0	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]
≥ 3	[redacted]	[redacted]	[redacted]	[redacted]
Number of Prior TNF Blocker Therapies, n (%)				
0	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]
≥ 3	[redacted]	[redacted]	[redacted]	[redacted]
Number of Failed Prior TNF Blocker Therapies,[a] n (%)				
0	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]
≥ 3	[redacted]	[redacted]	[redacted]	[redacted]
Number of Prior Systemic Therapies				
0	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]
≥ 3	[redacted]	[redacted]	[redacted]	[redacted]
Number of Failed Prior Systemic Therapies				
0	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]
≥ 3	[redacted]	[redacted]	[redacted]	[redacted]

BSA = body surface area; PASI = Psoriasis Activity and Severity Index; SD = standard deviation; TNF = tumour necrosis factor.
Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]
Note: Systemic therapies included conventional systemic (i.e., small molecule) and biologic therapies. Phototherapy included psoralen plus ultraviolet A light (PUVA) and ultraviolet B (UVB) light. Treatment failure information collected on the case report form was provided by the investigator and/or patient; treatment failure options included "never responded" or "lost response."

3.2.3 Interventions

Apremilast was administered orally, at a dose of 30 mg twice daily. The regimen in the placebo group was the same, and placebo tablets were described as being identical in appearance to apremilast.

Topical therapy within two weeks of randomization (including but not limited to topical corticosteroids, topical retinoid or vitamin D analogue preparations, tacrolimus, pimecrolimus, or anthralin/dithranol) was not allowed. Exceptions included low-potency corticosteroids (Class 6 or 7), which were allowed as background therapy for treatment of the face, axillae, and groin in accordance with the manufacturers' suggested usage during the course of the study. Patients with scalp psoriasis were permitted to use coal tar shampoo and/or salicylic acid scalp preparations on scalp lesions. An unmedicated skin moisturizer (e.g., Eucerin) was permitted for body lesions only. Patients should not have used these topical treatments within 24 hours prior to the clinic visit. Systemic therapy for psoriasis within four weeks prior to randomization (including but not limited to cyclosporine, corticosteroids, methotrexate, oral retinoids, mycophenolate, thioguanine, hydroxyurea, sirolimus, sulfasalazine, azathioprine, fumaric acid esters) was prohibited, as was the use of phototherapy within four weeks prior to randomization (i.e., ultraviolet light B [UVB], plus ultraviolet A light [PUVA]). Use of adalimumab, etanercept, efalizumab, infliximab, or certolizumab pegol was not allowed within 12 weeks prior to randomization, and use of alefacept, briakinumab, or ustekinumab was not allowed within 24 weeks prior to randomization.

3.2.4 Outcomes

The PASI score was determined for all patients throughout the study. PASI is a measure of psoriatic disease severity taking into account qualitative lesion characteristics (erythema, thickness, and scaling) and degree of skin surface area involvement on defined anatomical regions. PASI scores range from 0 to 72, with higher scores indicating greater disease severity. Erythema, thickness, and scaling are scored on a scale of 0 (none) to 4 (very severe) on four anatomic regions of the body: head, trunk, upper limbs, and lower limbs. Degree of involvement on each of the four anatomic regions is scored on a scale of 0 (no involvement) to 6 (90% to 100% involvement). The total qualitative score (sum of erythema, thickness, and scaling scores) is multiplied by the degree of involvement for each anatomic region and then multiplied by a constant. These values for each anatomic region are summed to yield the PASI score. The minimal clinically important difference (MCID) for PASI is unknown (see Appendix 5). PASI 75 represents a 75% reduction in PASI scores; PASI 50 is a 50% reduction, and so on.

sPGA is an assessment by the investigator of the overall disease severity at the time of evaluation. It is a five-point scale ranging from 0 (clear) to 4 (severe); the total score represents a summary assessment of the severity of the three primary signs of the disease: erythema, scaling, and plaque elevation. When making the assessment of overall severity, the investigator was instructed to factor in areas that had already been cleared (i.e., have scores of 0) and not just to evaluate remaining lesions for severity (i.e., the severity of each sign was to be averaged across all areas of involvement, including cleared lesions). In the event of different severities across disease signs, the sign that is the predominant feature of the disease was to be used to help determine the sPGA score. In addition to the above description provided in the protocol, investigators were provided guidance on the sPGA evaluation, which stipulated that if the outline of the original lesions could not be discerned, it was appropriate to conduct the evaluation in the context of the patient population (i.e., moderate to severe psoriasis with at least 10% BSA). The MCID for sPGA is unknown (see Appendix 5).
Affected BSA is a measurement of involved skin. The overall BSA affected by psoriasis is estimated based on the palm area of the patient's hand (entire palmar surface or "handprint"), which equates to approximately 1% of total BSA. The MCID for affected BSA is unknown (see Appendix 5).

The Nail Assessments/Nail Psoriasis Severity Index (NAPSI) assessed one target thumbnail or fingernail representing the worst nail psoriasis involvement at baseline for nail matrix psoriasis and nail bed psoriasis. For nail matrix, each quadrant of the nail is evaluated for any of the features of interest (pitting, leukonychia, red spots in the lunula, and crumbling), and scores are recorded based on the number of quadrants with any of these features (i.e., if all four quadrants have these features, the score would be 4). A similar protocol is followed for nail bed features (onycholysis, splinter hemorrhages, subungual hyperkeratosis, salmon patch dyschromia), again with scores from 0 to 4.[10] The sum of these two scores (nail matrix and nail bed) is the total score for that nail (therefore a range from 0 to 8). Thus higher scores indicate greater nail psoriasis severity. The number of fingers with psoriasis nail involvement was also counted. No MCID was found for this instrument.

The Scalp Physician Global Assessment (ScPGA) was used to assess scalp involvement if present at baseline. The six-point ScPGA scale ranges from 0 (clear) to 1 (minimal), 2 (mild), 3 (moderate), 4 (severe), and 5 (very severe). No MCID was found for this instrument.

The Palmoplantar Physician Global Assessment (PPPGA) was used to assess palms of hands and soles of feet for psoriasis involvement if present at baseline. The five-point PPPGA scale ranges from 0 (clear) to 1 (minimal), 2 (mild), 3 (moderate), and 4 (severe). No MCID was found for this instrument.

The Dermatology Life Quality Index (DLQI) was to be assessed by the patient upon arrival at the site before any other procedures or assessments were performed. The DLQI contains 10 items. The DLQI total score has a possible range of 0 to 30, with 30 corresponding to the worst quality of life, and 0 corresponding to the best score. The DLQI can be grouped into six subscales: symptoms and feelings, daily activities, leisure, work/school, personal relationships, and treatment. Scores for four of the subscales (symptoms and feelings, daily activities, leisure, and personal relationships) range from 0 to 6; scores for two of the subscales (work/school and treatment) range from 0 to 3. Higher scores indicate poorer quality of life. The MCID is considered to be 3.2 for the DLQI total score (see Appendix 5).

For pruritus, the patient was asked to assess itch in the previous week by placing a vertical stroke on a 100 mm visual analogue scale (VAS) on which the left-hand boundary represented no itch at all, and the right-hand boundary represented itch as worst itch imaginable. The distance from the mark to the left-hand boundary was to be recorded.

3.2.5 Statistical Analysis

a) ESTEEM-1 and ESTEEM-2

The primary efficacy end point was to be evaluated based on both the full analysis set (FAS) and the per-protocol (PP) analysis population. The analysis using the FAS was to be considered as the primary analysis.

Subgroup analyses for the primary end point and the major secondary efficacy end point based on baseline demographic (age, sex, race), baseline disease characteristics, as well as region were to be provided to determine the robustness of the treatment effect.

The proportions of patients in the two treatment groups (apremilast or placebo) who achieved at least a PASI 75 at week 16 in reference to the baseline visit were to be analyzed using a two-sided chi-square test at the 0.05 level. Missing values at week 16 were to be imputed using the last observation carried forward (LOCF) method, and supportive analyses were also performed using non-responder imputation

(NRI) (missing values were treated as non-responders) and by treating withdrawals due to adverse events (WDAEs) or lack of efficacy as non-responders, with other withdrawals imputed using LOCF.

A hierarchical approach was used to control for multiple statistical comparisons. The major secondary end point was the proportion of patients with an sPGA response, defined as an sPGA score of clear (0) or almost clear (1) with at least a two-point reduction from baseline at week 16. The two treatment groups (apremilast and placebo) were to be compared using a two-sided chi-square test at the 0.05 level, conditional on observing a statistically significant result for the primary analysis described in the previous section. Missing values at week 16 were to be imputed using the LOCF method for all secondary efficacy end points.

The other secondary end points were to be compared using the chi-square test for discrete variables, an analysis of covariance (ANCOVA) model for continuous variables, and the log-rank test for time-to-event variables. Statistical comparisons were to be conducted using two-sided tests at 0.05 level for the secondary end points in the order presented (per cent change in affected BSA, per cent change in PASI, proportion with PASI 50, change in pruritus VAS, and change in DLQI total). For any of the secondary end points, statistical significance was to be claimed only if statistically significant results were produced for the primary end point, the major secondary end point, and all other secondary efficacy end points before it. The changes or per cent changes from baseline between the two treatment groups (apremilast and placebo) were to be compared using an ANCOVA model with treatment as the factor and the baseline score as the covariate.

Descriptive statistics were to be presented for exploratory end points. Specifically, for continuous variables, summary statistics (number of patients with event [n], mean, standard deviation [SD], median, minimum, and maximum) for baseline, specified time points and changes from baseline were to be provided. Categorical variables were to be summarized with frequency tabulations. The Kaplan-Meier procedure was to be used to characterize the time to achieve at least PASI 50 or PASI 75 during the placebo-controlled phase.

In ESTEEM-1, the proposed sample size was to provide a sufficient safety database. Approximately 825 patients were to be randomized into the study, with 550 patients in apremilast group and 275 patients in the placebo group. After week 16, all patients were treated with apremilast until week 32. If the event frequency was ≥ 0.2% with apremilast treatment, the chance to observe at least one event with 825 patients was > 80%. In ESTEEM-2, sample size estimation for the primary end point was based on results of a phase 2b study. A chi-square test with a 0.05 two-sided significance level would provide 90% power to detect a 20% difference (30% versus 10%) between apremilast and placebo for the proportion of patients achieving at least PASI 75 at week 16 when the total sample size was approximately 189 with a 2:1 randomization.

b) Analysis populations

For both studies, the safety population analysis set was to consist of all patients who were randomized and received at least one dose of study drug.

The FAS was to consist of all patients who were randomized as specified in the protocol. The PP population analysis set was to consist of all patients included in the safety population who had at least one post-treatment PASI evaluation and no protocol violations.

3.3 Patient Disposition

In ESTEEM-1, 12% of patients discontinued from the placebo-controlled phase, with similar proportions between groups, while in ESTEEM-2, 13% of apremilast patients and 18% of placebo patients discontinued. The most common reason for discontinuation in both studies was AE. The vast majority of patients who completed the placebo-controlled phase entered the maintenance phase.

TABLE 6: PATIENT DISPOSITION

	ESTEEM-1		ESTEEM-2	
	Apremilast	**Placebo**	**Apremilast**	**Placebo**
Screened, N	1,121		569	
Randomized, N (%)	562	282	275	138
Randomized and Treated, N (%)			272 (99)	136 (99)
Discontinued, N (%)	59 (11)	33 (12)	35 (13)	25 (18)
Adverse event	23 (4)	5 (2)	12 (4)	8 (6)
Lack of efficacy	2 (< 1)	7 (3)	3 (1)	2 (1)
Non-compliance with study drug	7 (1)	0	1 (< 1)	0
Withdrawal by patient	12 (2)	9 (3)	9 (3)	7 (5)
Death	0	1 (< 1)		
Lost to follow-up	7 (1)	9 (3)	6 (2)	6 (4)
Protocol violation	7 (1)	1 (< 1)	2 (1)	1 (1)
Other	1 (< 1)	1 (< 1)	2 (1)	1 (1)
Completed and Entered Maintenance Phase	494 (88)	245 (87)	234 (85)	108 (78)
Completed But Did Not Enter Maintenance Phase	9 (2)	4 (1)	5 (2)	4 (3)
Primary Reason for Discontinuation				
Adverse event	[redacted]	[redacted]	[redacted]	[redacted]
Lack of efficacy	[redacted]	[redacted]	[redacted]	[redacted]
Non-compliance with study drug	[redacted]	[redacted]	[redacted]	[redacted]
Withdrawal by patient	[redacted]	[redacted]	[redacted]	[redacted]
Death	[redacted]	[redacted]	[redacted]	[redacted]
Lost to follow-up	[redacted]	[redacted]	[redacted]	[redacted]
Study terminated by sponsor	[redacted]	[redacted]	[redacted]	[redacted]
Protocol violation	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]
Entered Post-treatment Observational Follow-up Phase	[redacted]	[redacted]	[redacted]	[redacted]
Completed	[redacted]	[redacted]	[redacted]	[redacted]
Full Analysis Set, N	562 (100)	282 (100)	274 (100)	137 (99)
Per-Protocol, N	555 (99)	276 (98)	266 (97)	134 (97)
Safety, N	560 (100)	282 (100)	272 (99)	136 (99)

Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

3.4 Exposure to Study Treatments

In ESTEEM-1 and ESTEEM-2, exposure was similar between apremilast and placebo groups (ESTEEM-1: 15.0 versus 14.8 weeks; ESTEEM-2: 14.6 versus 14.0 weeks).

3.5 Critical Appraisal

3.5.1 Internal Validity

Randomization was performed using an interactive voice response system (IVRS), and appropriate measures appear to have been taken to ensure allocation concealment. Both studies were double-blinded, and the placebo intervention was described by the manufacturer as matched in appearance to apremilast. Apremilast is associated with gastrointestinal adverse effects, and it is possible that this may have led to unblinding in some instances, as patients who experienced these adverse effects may have speculated that they were on apremilast. It is not known what the precise impact of unblinding would be on investigator-assessed outcomes such as PASI, and the effectiveness of the blinding does not appear to have been tested.

The manufacturer accounted for multiplicity in its secondary outcomes by using a hierarchical testing procedure, which is considered an acceptable strategy for controlling for multiple comparisons. In this procedure, statistical testing is only continued on subsequent outcomes for as long as testing reveals statistical significance on the previous outcome. There were also exploratory outcomes that were tested that fell outside of the hierarchy, although the only one that was a key efficacy outcome in this review was the EuroQol 5-Dimensions Health-Related Quality of Life Questionnaire (EQ-5D). Because it fell outside of the hierarchy, the EQ-5D results must be interpreted with caution.

In ESTEEM-1, the discontinuation rate was relatively low and similar between apremilast (11% of patients) and placebo (12%). However, in ESTEEM-2 there was a slightly higher proportion of withdrawals in the placebo group (18% of patients) versus apremilast (13%). It is not clear why the discontinuation rate was higher in the placebo group of this study compared with the other groups, including ESTEEM-1. The two most common reasons for discontinuation in the placebo group were AEs and “withdrawal by patient”; the former is typically a common reason for study withdrawal, while the latter is a general description that lacks sufficient detail to appraise. A higher discontinuation rate in the placebo group might introduce bias depending on how missing data are accounted for. For the primary outcome — PASI 75 — missing data were accounted for by using LOCF. In the placebo group, where PASI 75 responses were low, this is perhaps unlikely to introduce bias in either direction, as there appeared to be little impact on PASI 75 in the placebo group.

3.5.2 External validity

According to the clinical expert, the baseline characteristics of the populations enrolled in ESTEEM-1 and ESTEEM-2 appear consistent with that of the population in which apremilast would be used. Both studies were multinational with Canadian sites, further suggesting the results may be generalizable to the Canadian population.

The placebo-controlled phase of each included study was only 16 weeks in duration. Given that apremilast employs a novel mechanism of action, this might not be sufficient follow-up to assess efficacy and more notably, safety. For example, the TNF inhibitors have, throughout their history, been associated with elevated risk of certain cancers. Because risk of cancer can be determined only over the course of several years follow-up, the association between TNF inhibitors and cancer took many years to be elucidated, and to this day remains unclear for certain cancers. Therefore, while a 16-week follow-up may be sufficient to determine certain safety issues, it is not likely long enough for others. Psoriasis is a chronic condition, and patients may be treated with a single therapy for many years.

There is a lack of direct, head-to-head comparisons of apremilast versus another active control, such as methotrexate or one of the biologics. Although there is an ongoing trial of apremilast compared with

etanercept, this study is not due to be completed until August 2015. Therefore, only indirect comparisons can be made between apremilast and these other drugs, and these comparisons have significant limitations that introduce bias into the analysis and reduce confidence in any conclusions that are drawn.

The primary outcome of both included studies was the proportion of patients achieving a PASI 75, and this appears consistent with the other studies in chronic plaque psoriasis. Some of the limitations of PASI as an assessment include the fact that it often does not correlate well with quality of life. In the included studies, this does not appear to have been an issue, at least for the DLQI, as statistically significant improvements were seen across PASI and DLQI.

3.6 Efficacy

Only those efficacy outcomes identified in the review protocol are reported below (Section 2.2, Table 3). See APPENDIX 4: DETAILED OUTCOME DATA for detailed efficacy data. Data for the maintenance and re-randomization phases, which did not meet inclusion criteria, are summarized in Appendix 4 (Table 11, Table 12, Table 13, and Table 14).

3.6.1 Psoriasis Area and Severity Index

The proportion of patients achieving PASI 75 at 16 weeks was the primary outcome of both ESTEEM-1 and ESTEEM-2 (Table 7). In ESTEEM-1, 33% of apremilast patients and 5% of placebo patients achieved PASI 75 after 16 weeks, a statistically significant difference between groups (difference in proportions of 27.8% [95% CI, 23.1 to 32.5], $P < 0.0001$). In ESTEEM-2, 29% of apremilast patients and 6% of placebo patients achieved PASI 75 after 16 weeks, and this difference was also statistically significant (difference in proportions of 23.0% [95% CI, 16.3% to 29.6%], $P < 0.0001$).

PASI 50 after 16 weeks was achieved by a larger proportion of apremilast versus placebo patients in both ESTEEM-1 (59% versus 17%, difference in proportions of 41.7% [95% CI, 35.7 to 47.7], $P < 0.0001$) and in ESTEEM-2 (56% versus 20%, difference in proportions of 35.8% [95% CI, 26.9 to 44.7], $P < 0.0001$).

There was a statistically significant reduction from baseline to 16 weeks in the least squares (LS) mean (95% CI) PASI scores in each of ESTEEM-1 (–35.3 [–39.9 to –30.6], $P < 0.0001$) and ESTEEM-2 (–34.8 [–42.4 to –27.2], $P < 0.0001$).

a) Subgroups

Subgroup data were reported for both ESTEEM-1 and ESTEEM-2 (Table 10). In ESTEEM-1, of the patients with a history of psoriatic arthritis, [redacted]% of apremilast patients and [redacted]% of placebo patients had a PASI 75 after 16 weeks, while the proportions were [redacted]% versus [redacted]%, respectively, in patients without a history of psoriatic arthritis. In ESTEEM-2 after 16 weeks, [redacted]% of apremilast patients versus [redacted]% of placebo patients with a history of psoriatic arthritis achieved PASI 75, while in patients without a history of psoriatic arthritis the proportions were [redacted]% and [redacted]%, respectively.

In ESTEEM-1, in patients who failed prior phototherapy, the proportion of responders was [redacted]% with apremilast versus [redacted]% with placebo after 16 weeks and in patients who had not failed prior phototherapy, the proportions were [redacted]% and [redacted], respectively. Prior treatment failure could be reported by either the investigator or the patient, and failure could be defined as either having never responded or loss of response. In ESTEEM-2, the proportions in patients who had failed prior phototherapy were [redacted]% versus [redacted]%, and in those who had not failed prior phototherapy, the proportions were [redacted]% versus [redacted]%, respectively.

In ESTEEM-1, in patients who failed prior systemic therapy, the proportion of responders was [redacted]% with apremilast versus [redacted]% with placebo, and in patients who had not failed prior systemic therapy, the proportions were [redacted]% and [redacted]%, respectively. In ESTEEM-2, the proportions in patients who had failed prior systemic therapy were [redacted]% versus [redacted]%, and in those who had not failed prior systemic therapy, the proportions were [redacted]% versus [redacted]%, respectively.

In ESTEEM-1, in patients who failed prior biologic therapy, the proportion of responders was [redacted]% with apremilast versus [redacted]% with placebo, and in patients who had not failed prior biologic therapy, the proportions were [redacted]% and [redacted]%, respectively. In ESTEEM-2, the proportions of patients who had failed prior biologic therapy were [redacted]% versus [redacted]%, and in those who had not failed prior biologic therapy, the proportions were [redacted]% versus [redacted]%, respectively.

In ESTEEM-1, in patients who failed prior anti-TNF therapy, the proportion of responders was [redacted]% with apremilast versus [redacted]% with placebo, and in patients who had not failed prior anti-TNF therapy, the proportions were [redacted]% and [redacted]%, respectively. In ESTEEM-2, the proportions of patients who had failed prior biologic therapy were [redacted]% versus [redacted]%, and in those who had not failed prior biologic therapy, the proportions were [redacted]% versus [redacted]%, respectively.

3.6.2 Quality of Life

There was a statistically significant reduction (improvement) from baseline to 16 weeks in DLQI total scores for apremilast versus placebo in ESTEEM-1 (difference in LS means of –4.5 [95% CI, –5.4 to –3.6], $P < 0.0001$), and in ESTEEM-2 (difference in LS means of –4.0 [95% CI, –5.3 to –2.8], $P < 0.0001$). The MCID for the DLQI is estimated to be 3.2; therefore, these differences versus placebo are also clinically significant (Table 7).

Quality of life was also assessed using the EQ-5D; however, this was an exploratory outcome and not part of the statistical hierarchy, and any statistical analysis should therefore be interpreted with caution. There was a statistically significant improvement in index scores from baseline for apremilast over placebo in both studies, ESTEEM-1 (difference in LS means [two-sided 95% CI]: [redacted], *P* [redacted]) and ESTEEM-2 (difference in LS means [two-sided 95% CI]: [redacted], *P* = [redacted]) (Table 7). However, statistically significant improvements on the EQ-5D VAS were only seen in ESTEEM-2 (difference in LS means [2-sided 95% CI]: [redacted], *P* = [redacted]) and not in ESTEEM-1 (difference in LS means [two-sided 95% CI]: [redacted], *P* = [redacted]). The MCID for index scores in psoriasis is considered to be between 0.09 and 0.20, suggesting that these improvements in EQ-5D index scores are unlikely to be clinically significant in this population.

3.6.3 Physician Global Assessment

A larger proportion of apremilast patients achieved an sPGA response after 16 weeks compared with placebo in ESTEEM-1: 22% versus 4%, and this difference was statistically significant (difference in proportions of 17.8% [95% CI, 13.7 to 21.9], $P < 0.0001$) (Table 7). Similarly, in ESTEEM-2, apremilast had a higher proportion of sPGA responders than placebo after 16 weeks: 20% versus 4% (difference in proportions of 16.1% [95% CI, 10.2 to 21.9], $P < 0.0001$).

3.6.4 Affected Body Surface Area

In ESTEEM-1, the LS mean (95% CI) change from baseline to week 16 in percentage-affected BSA was –47.8% (–51.0 to –44.6) with apremilast and –7.0% (–11.5 to –2.4) with placebo, and this difference was statistically significant (difference in LS means of –40.78% [–46.34 to –35.21], $P < 0.0001$) (Table 7). In ESTEEM-2, the LS mean (95% CI) change from baseline to week 16 in percentage-affected BSA was –48.4% (–53.6 to –43.2) with apremilast and –6.3% (–13.5 to 1.1) with placebo, and this difference was

statistically significant (difference in LS means of –42.15% [–51.11 to –33.20], *P* < 0.0001). The MCID is unknown for affected BSA.

3.6.5 Other Efficacy Outcomes

Other efficacy outcomes included symptom measures, including the pruritus VAS (Table 9). In ESTEEM-1 (difference in LS means [two-sided 95% CI] of –24.2 [–28.7 to –19.8], *P* < 0.0001) and ESTEEM-2 (–21.3 [–28.4 to –14.2], *P* < 0.0001) there was a statistically significant improvement in pruritus for apremilast versus placebo.

Nail symptoms were also assessed using NAPSI. NAPSI scores improved from baseline to week 16 for apremilast versus placebo in both ESTEEM-1 (difference in LS means [two-sided 95% CI] –29.1 [–28.6 to –16.5], *P* < 0.0001) and ESTEEM-2 (difference in LS means [two-sided 95% CI] –22.9 [–38.9 to –6.9], *P* = 0.0052) (Table 9).

TABLE 7: KEY EFFICACY OUTCOMES

	ESTEEM-1		ESTEEM-2	
PASI 75	**Apremilast N = 562**	**Placebo N = 282**	**Apremilast N = 274**	**Placebo N = 137**
Patients at week 16, n (%) — FAS/LOCF	186 (33)	15 (5)	79 (29)	8 (6)
Difference in proportions (95% CI)[a]	27.8 (23.1 to 32.5)		23.0 (16.3 to 29.6)	
P value	*P* < 0.0001		*P* < 0.0001	
NRI[b]/LOCF	183 (33)	14 (5)	79 (29)	7 (5)
sPGA				
sPGA response, week 16, n (%)[c]	122 (22)	11 (4)	56 (20)	6 (4)
Difference in proportions (95% CI)[a]	17.8 (13.7 to 21.9)		16.1 (10.2 to 21.9)	
P value	*P* < 0.0001		*P* < 0.0001	
Affected BSA				
Mean (SD) at baseline	24.4 (14.7)	25.0 (14.3)	25.5 (15.5)	27.6 (15.9)
Mean (SD) % change at week 16	–47.8 (38.5)	–6.9 (38.9)	–48.5 (40.8)	–6.1 (47.6)
LS mean [95% CI]	–47.8 [–51.0 to –44.6]	–7.0 [–11.5 to –2.4]	-48.4 [53.6 to –43.2]	–6.3 [–13.5 to 1.1]
Difference in LS means (2-sided 95% CI)[d]	–40.78 (–46.34 to –35.21)		–42.15 (–51.11 to –33.20)	
P value	*P* < 0.0001		*P* < 0.0001	
Change in PASI				
Mean (SD) at baseline	18.7 (7.2)	19.3 (7.4)	19.0 (7.1)	20.1 (8.0)
Mean (SD) % change at week 16	–52.1 (32.8)	–16.7 (31.5)	–50.9 (34.0)	–15.8 (41.3)
LS mean (95% CI)	–52.1 (54.7 to –49.4)	–16.8 (–20.6 to –13.0)	–50.8 (–55.2 to –46.4)	–16.0 (–22.2 to –9.8)
Difference in LS means (2-sided 95% CI)[d]	–35.3 (–39.9 to –30.6)		–34.8 (–42.4 to –27.2)	
P value	*P* < 0.0001		*P* < 0.0001	
DLQI Total Score				
Mean (SD) at baseline	12.7 (7.1)	12.1 (6.7)	12.5 (7.1)	12.8 (7.1)
Mean (SD) change at week 16	–6.6 (–6.7)	–2.1 (5.7)	–6.7 (7.0)	–2.8 (7.2)
LS mean change (95% CI)	–6.6 (–7.1 to –6.1)	–2.1 (–2.8 to –1.3)	–6.7 (–7.5 to –6.0)	–2.7 (–3.7 to –1.7)
Difference in LS means (2-sided 95% CI)	–4.5 (–5.4 to –3.6) [e]		–4.0 (–5.3 to –2.8) [f]	

	ESTEEM-1		ESTEEM-2	
P value	*P* < 0.0001		*P* < 0.0001	
EQ-5D VAS				
Mean (SD) at baseline	[redacted] N=[redacted]	[redacted] N=[redacted]	[redacted] N=[redacted]	[redacted] N=[redacted]
Mean (SD) change at week 16	[redacted]	[redacted]	[redacted]	[redacted]
LS mean change from baseline (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Difference in LS means (2-sided 95% CI)[f]	[redacted]		[redacted]	
2-sided *P* value	*P* = [redacted]		*P* = [redacted]	
EQ-5D Index Value Score				
Mean (SD) at baseline	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD) change at week 16	[redacted]	[redacted]	[redacted]	[redacted]
LS mean change from baseline (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Difference in LS means (2-sided 95% CI)[g]	[redacted]		[redacted]	
2-sided *P* value	*P* < [redacted]		*P* = [redacted][h]	

ANCOVA = analysis of covariance; ANOVA = analysis of variance; BSA = body surface area; CI = confidence interval; DLQI = Dermatology Life Quality Index; EQ-5D = EuroQol 5-Dimensions Health-Related Quality of Life Questionnaire; FAS = full analysis set; LOCF = last observation carried forward; LS = least squares; NRI = non-responder imputation; PASI = Psoriasis Area and Severity Index; SD = standard deviation; sPGA = static Physician Global Assessment; VAS = visual analogue scale.

[a] Two-sided 95% CI is based on the normal approximation. Two-sided *P* value is based on the two-sided chi-square test.

[b] This is a sensitivity analysis using both LOCF and NRI to account for missing values. Non-responder imputations were patients who discontinued before week 16 or whose PASI evaluations were missing.

[c] sPGA response: score of 0 or 1 with ≥ two-point reduction from baseline.

[d] Based on an ANCOVA model for the per cent change from baseline at week 16, with treatment group as a factor and the baseline value as a covariate. Means (LS means), difference in LS means, and *P* values were adjusted by covariate. The two-sided *P* value for slope homogeneity is > 0.05.

[e] Based on an ANOVA model for the change from baseline at week 16, with treatment group as a factor (an ANOVA model). Unadjusted means and *P* values are provided. The two-sided *P* value for slope homogeneity is < 0.05.

[f] Based on an ANCOVA model for the per cent change from baseline at week 16, with treatment group as a factor and the baseline value as a covariate. Means (LS means) and *P* values were adjusted by covariate. The two-sided *P* value for slope homogeneity is > 0.05.

[g] Based on an ANCOVA model for the change from baseline at week 16. The model includes treatment group as a factor (analysis of variance model). The unadjusted means and *P* value are provided. The two-sided *P* value for slope homogeneity is < 0.05.

[h] *P* values ≤ 0.050 and considered nominally significant, as there was no adjustment for multiplicity based on hierarchical testing.

Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

3.6.6 Harms

Only those harms identified in the review protocol are reported below (see 2.2.1, Protocol). See APPENDIX 4: DETAILED OUTCOME DATA for detailed harms data.

3.6.7 Adverse Events

In ESTEEM-1, 69% of apremilast patients and 56% of placebo patients reported an AE after 16 weeks of therapy, while in ESTEEM-2, 68% of apremilast patients and 60% of placebo patients experienced an AE (Table 8). The most common AEs were diarrhea (18% of apremilast patients versus 7% of placebo patients across studies) and nausea (17% of apremilast patients and 7% of placebo patients).

3.6.8 Serious Adverse Events

Serious adverse events (SAEs) were reported in 2% of apremilast patients and 3% of placebo patients after 16 weeks in ESTEEM-1, and in 2% of each of the apremilast and placebo groups in ESTEEM-2 (Table 8). No single SAE occurred in more than a single patient.

3.6.9 Withdrawals due to Adverse Events

WDAEs occurred in 5% of apremilast patients and 3% of placebo patients after 16 weeks in ESTEEM-1, and in ESTEEM-2, 6% and 5% of patients, respectively (Table 8). The most common reason for withdrawal across groups was nausea.

3.6.10 Mortality

In ESTEEM-1, there was one death in each of the apremilast and placebo groups. There were no deaths in ESTEEM-2 (Table 8Table 8).

3.6.11 Notable Harms

See notes for AEs (section 3.6.7). Aside from gastrointestinal AEs, weight loss was a notable harm for this review. Manufacturer-provided data indicated that in both ESTEEM studies, weight loss as an AE was reported in 1.4% (17/1,184) of patients treated with apremilast (weeks 0 to 52) and in 0.2% (1/418) of patients treated with placebo (weeks 0 to 16). Of note, however, the proportion of patients with clinically significant weight loss (i.e., weight decreases of > 5% of baseline body weight), not specified as an AE, was [redacted]% for apremilast versus [redacted]% for placebo in ESTEEM-1, and [redacted]% versus [redacted]%, respectively, in ESTEEM-2.

TABLE 8: HARMS

	ESTEEM-1		ESTEEM-2	
AEs	**Apremilast N = 562**	**Placebo N = 282**	**Apremilast N = 272**	**Placebo N = 136**
Patients with > 0 AEs, n (%)	388 (69)	157 (56)	185 (68)	82 (60)
Most common AEs				
Diarrhea	105 (19)	20 (7)	43 (16)	8 (6)
Nausea	88 (16)	19 (7)	50 (18)	9 (7)
Upper respiratory tract infection	57 (10)	21 (7)	13 (5)	6 (4)
SAEs				
Patients with > 0 SAEs, n (%)	12 (2)	8 (3)	5 (2)	3 (2)
Most common SAEs				
None in more than one patient	–	–	–	–
WDAEs				
WDAEs, n (%)	29 (5)	9 (3)	15 (6)	7 (5)
Most common reasons				
Nausea	10 (2)	1 (< 1)	3 (1)	0
Diarrhea	7 (1)	1 (< 1)	NR	NR
Psoriasis	–	–	2 (1)	3 (2)
DEATHS				
Number of deaths, n (%)	1 (< 1)	1 (< 1)	0	0
Most common reasons				
Reason	Cardiac failure	Suicide		
NOTABLE HARMS				
Weight loss reported as an AE	NR	NR	NR	NR
Weight loss > 5% of baseline	[redacted]	[redacted]	[redacted]	[redacted]

AE = adverse event; NR = not reported; SAE = serious adverse event; WDAE = withdrawal due to adverse event.
Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

4. DISCUSSION

4.1 Summary of Available Evidence

Two pivotal, phase 3, placebo-controlled DB RCTs — ESTEEM-1 and ESTEEM-2 — met the inclusion criteria for this review. ESTEEM-1 (N = 844) and ESTEEM-2 (N = 413) both randomized patients with moderate to severe chronic plaque psoriasis in a 2:1 manner to either apremilast 30 mg twice daily or placebo, over an initial RCT phase of 16 weeks. This 16-week phase was followed by a 20-week maintenance phase where placebo patients were switched to apremilast 30 mg twice daily, and then a re-randomization phase where responders were re-randomized to apremilast or placebo. The initial 16-week placebo-controlled phase was the focus of this review, as it met the key inclusion criteria of an RCT design. The primary outcome of both studies was proportion of patients achieving PASI 75, and apremilast was statistically significantly superior to placebo for this outcome in both studies. More apremilast patients had an sPGA response compared with placebo, and the proportion of affected BSA was statistically significantly reduced with apremilast versus placebo. MCIDs have not been established for these outcomes; however, according to the clinical expert consulted for the review, these results should be considered clinically significant. Quality of life was also statistically and clinically significantly improved versus placebo using the DLQI in both studies. Using a more generic instrument, the EQ-5D, quality of life was statistically significantly improved versus placebo in ESTEEM-2 but not in ESTEEM-1 on the VAS, although index scores were statistically significantly improved versus placebo in both studies. The EQ-5D findings were an exploratory analysis; therefore, no adjustment was made for multiplicity, thus these results are potentially susceptible to inflated type 1 error. Given the range of MCIDs for EQ-5D index scores in psoriasis, these differences are unlikely to be clinically significant. Gastrointestinal AEs such as nausea and diarrhea were the most common AEs, and were identified a priori for this review as notable harms with this drug. Gastrointestinal AEs also appeared to be the most common reason for withdrawal, suggesting these may impact adherence and persistence with therapy. There were no specific SAEs that occurred in more than one patient, and thus no clear indication of any specific issues regarding serious harms.

4.2 Interpretation of Results

4.2.1 Efficacy

Data from the dermatology-focused DLQI scale indicate that patients taking apremilast experience a statistically and clinically significant improvement in their quality of life over the course of 16 weeks when compared with placebo. In addition to the dermatological symptoms such as pruritus, psoriasis lesions have an impact on physical appearance, and the patient input submitted to CDR does indeed suggest that psoriasis negatively affects quality of life. Findings for the generic EQ-5D instrument were less clear, indicating an inconsistent improvement in quality of life using the EQ-5D VAS score, and a lack of clinically significant improvement in EQ-5D index scores, based on the MCID for psoriasis. It is not clear why there were inconsistent findings for EQ-5D VAS scores, as there were no obvious differences in baseline characteristics between studies. However, the DLQI is specific to dermatology and has been validated; therefore, findings from this instrument are more likely to reflect changes in quality of life for these patients with respect to treatment. Additionally, the EQ-5D was an exploratory outcome in the ESTEEM-1 and ESTEEM-2 trials; thus, any statistical analysis should be interpreted with caution. An important question is whether these improvements in quality of life on the DLQI persist over time, as all measures of patient-reported outcomes beyond the initial 16-week DB phase would potentially be biased by the lack of blinding. In their input to CDR, patients emphasized their frustration with the fact that the beneficial effects of many therapies for psoriasis dissipate with time, and with only a 16-week

follow-up for the DB phase, it is not clear whether this will be the case with apremilast with respect to quality of life.

As noted above, patients expressed frustration and concern with the loss of response over time exhibited by other therapies for psoriasis. The ESTEEM-1 and ESTEEM-2 studies only had a 16-week DB phase; however, there were two additional phases beyond this, followed by an extension, that might assist in determining whether responses persist for some of the more objective outcomes, such as PASI. Data from the maintenance phase (weeks 16 through 32), where all patients received apremilast, suggest that the effects of apremilast persisted with the additional 16 weeks of therapy in the apremilast group, while the group that switched from placebo to apremilast had similar responses to those receiving 32 weeks of apremilast (Table 11). These findings are limited, however, by the fact that at this point in the study, all patients would know that they were on apremilast; therefore, blinding is lost. In the re-randomization phase, more patients who were switched to placebo lost their PASI 75 responses compared with those who continued on apremilast; however, this is not a surprising finding (Table 12).

Subgroup analysis was presented for patients who had failed prior therapies, systemic therapies, phototherapies, and biologics. Treatment failure, according to the manufacturer, was defined as patients who either never responded to the drug or lost response to the drug, and it appears the determination of failure could be made by investigator and/or patient. However, there were a number of patients (approximately one-third of the randomized population) who were identified as having never tried prior systemic therapies at baseline. Given the proposed manufacturer's listing criteria, which state that patients are to have had an inadequate response to, be intolerant of, or have a contraindication to conventional systemic therapy, this seems to be a large number of patients who would never have tried these therapies previously. It is possible that this one-third of patients who were naive to systemic therapy represents patients with a contraindication to therapy; however, no further details were provided as to why these patients had never received therapy. Prior failure on systemic therapy did not appear to have any impact on PASI 75 response to apremilast versus placebo in ESTEEM-1; however, in ESTEEM-2, the placebo response appeared to be higher in patients who had failed prior conventional systemic therapy. Additionally, the fact that prior treatment failure could potentially be determined by the patient introduces uncertainty into whether all patients who were classified as prior treatment failures were in fact treatment failures. Moreover, the subgroup specified in the manufacturer's listing request (i.e., those who had an inadequate response to, or are intolerant or contraindicated to a conventional systemic therapy) was not a pre-specified subgroup in the ESTEEM studies. The manufacturer provided additional subgroup analyses specific to the listing criteria, with statistical analyses, when submitting their comments on this CDR report.[11] However, these were post hoc analyses and are thus hypothesis-generating. Given this analysis and the limitations outlined above, there is a high degree of uncertainty related to the efficacy and harms of apremilast in the population proposed under the manufacturer's listing criteria.

The manufacturer provided a network meta-analysis; however, no statistical testing was performed that directly compared apremilast to other active comparators, and instead all comparisons were made to placebo. Results suggest that the efficacy of small molecule inhibitors such as apremilast is less than that of the biologics, with respect to PASI responses, although these data can be considered only hypothesis-generating as no formal statistical analysis was reported. Another major limitation of the analysis is a lack of harms data. The biologics have had longstanding issues with respect to increased risk of infection as well as issues with long-term use, such as cancer, while the existing small molecules have significant safety issues associated with their use. Although there do not appear to be any significant safety issues

with apremilast, a formal indirect comparison of harms data would have provided more insight into the potential for an improved harms profile with apremilast.

4.2.2 Harms

In patients' input submitted to CDR, it is clear that adverse effects are a significant concern for patients being treated for psoriasis. Patients are concerned about the well-known adverse effects associated with systemic therapies, and seem particularly concerned with the various harms associated with the small molecule inhibitors such as cyclosporine and methotrexate. The most common AEs with apremilast were gastrointestinal in nature (nausea and diarrhea), and these are also noted in the product monograph. There was some evidence that these adverse effects may impact persistence with therapy, as they were also the most common reason for WDAEs. These adverse effects have been seen with other phosphodiesterase inhibitors, so this may be a class effect. There were no notable differences in risk of serious gastrointestinal AEs with apremilast, and in fact there were no serious AEs that occurred in more than one patient. Therefore, in the short term, apremilast does not appear to have any notable safety issues, although there may be some issues with persistence due to gastrointestinal events.

The long-term safety of apremilast is less clear. The placebo-controlled phase of both ESTEEM-1 and ESTEEM-2 lasted only 16 weeks, and this is not likely of sufficient duration to reveal any long-term safety issues, if any exist. Longer-term safety data were available from the two subsequent phases, as well as the extension; however, the two phases are no longer blinded, and the extension lacks a control group (see Appendix 6). The systemic small molecule inhibitors used in psoriasis (methotrexate and cyclosporine) have well-established safety issues, both in the short and long term. Similarly, the biologics also have short- and longer-term safety issues, including cancer. As noted above, there do not appear to be any clear efficacy advantages of using apremilast versus these other comparators, thus it is important that any potential safety advantages be fully characterized.

5. CONCLUSIONS

Two multinational, manufacturer-sponsored pivotal phase 3, DB RCTs — ESTEEM-1 and ESTEEM-2 — met the inclusion criteria for this review. Both studies, ESTEEM-1 (N = 844) and ESTEEM-2 (N = 413) randomized patients with moderate to severe chronic plaque psoriasis in a 2:1 manner to either apremilast 30 mg twice daily or placebo, over a 16-week DB comparison period. Apremilast was statistically significantly superior to placebo after 16 weeks for proportion of patients achieving a PASI 75, the primary outcome of both studies. Apremilast also improved sPGA responses versus placebo, and reduced the affected BSA versus placebo, all statistically significant differences. Quality of life was statistically and clinically significantly improved after 16 weeks versus placebo in both studies using a dermatology-specific instrument, the DLQI. The most common AEs with apremilast were gastrointestinal-related (nausea and diarrhea), and these were also the most common reasons for WDAEs. There were no clear indications of any serious harms issues with apremilast, although interpretation of this finding is limited somewhat by the relatively short 16-week follow-up in the DB comparative phase. A manufacturer-submitted NMA suggested the biologics provided superior efficacy to the small molecule drugs for psoriasis, including apremilast, based on rank ordering alone. There is uncertainty with respect to the conclusions from the NMA because the manufacturer did not provide any statistical indirect comparison estimates.

APPENDIX 1: PATIENT INPUT SUMMARY

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

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

The Canadian Skin Patient Alliance (CSPA), working with and supported by the Canadian Association of Psoriasis Patients (CAPP), submitted the input for this review. CSPA is a non-profit organization serving patients with dermatological conditions, and focuses on advocacy, education, and support for more than 20 allied or affiliated disease-specific organizations. CAPP is a non-profit organization supporting patients with psoriasis and psoriatic arthritis.

CSPA reported receiving support from the following sources: AbbVie, Amgen, Celgene, GlaxoSmithKline, Leo Pharma, Merck, Novartis, Roche, and Valeant. CAPP reported receiving support from the following sources: AbbVie, Amgen, Janssen, Leo Pharma, Celgene, and Pfizer. The submission was prepared by CSPA and CAPP staff and declared that no outside corporate source influenced the submission.

2. Condition and Current Therapy-Related Information

Information for this submission was obtained from patient questionnaires (administered online through social media, or sent to research physicians who then passed it on to patients), patient feedback from disease information sessions, and the Multinational Assessment of Psoriasis and Psoriatic Arthritis (MAPP) study.

Persons with psoriasis experience lesions and plaques on their body. The physical symptoms of their disease include painful, bleeding, cracking, crusting, and flaking lesions and plaques, and many experience severe itch. Many patients report losing sleep because of itch, and some have put their skin in vinegar or scratched themselves raw as “pain is preferable to itch.” Patients also report limited activities due to their psoriasis: inability to participate in sports, limited movement due to joint pain, inability to perform day-to-day tasks, fear of job loss, missed days of work due to symptoms, and concentration issues related to sleep loss.

While body surface covered by plaques can be an indication of severity, many patients find lesion location to be as important. Patients report lesions in sensitive areas that affect perception of attractiveness and sexuality. Psychosocially, they experience stigma, depression, suicidality, shame, feelings of helplessness, and frustration, and they fear shunning from others. One patient reported being asked to leave his gym because other patrons were uncomfortable with psoriasis. Understandably, then, persons with psoriasis may also isolate themselves from others, affecting relationships with family and friends. Additionally, patients may experience comorbid conditions, such as diabetes, depression-related weight gain, and heart disease, while struggling with managing their psoriasis.

Psoriasis also affects the lives of a patient’s caregivers and loved ones. The depression and pain of psoriasis patients affects the entire family. Constant cleaning associated with the flaking of skin is a burden, and additional help may be required to clean and manage household tasks. Patients may avoid or miss important events, such as school functions and outings. Children of parents with psoriasis live in fear of also getting the disease.

In general, patients will respond to a therapy for a time but then experience “biologic fatigue,” where the treatment loses efficacy. This was identified as a major factor in the management of psoriasis.

Patients want options for therapy, as their current treatment regimens may loss effectiveness in the management of their psoriasis. Current treatment includes methotrexate, cyclosporine, etanercept, adalimumab, infliximab, ustekinumab, and phototherapy. Adverse effects include treatment toxicity, such as liver and kidney damage or the fear of liver and kidney damage; nausea; headaches; and feelings of malaise, especially for methotrexate and cyclosporine. Additionally, the high cost of therapy, multiple injections, lack of therapy options, and limited ability to travel are limitations felt by patients. Also noted was the struggle to access therapies as patients and their doctors must repeatedly file paperwork to qualify for treatments.

3. Related Information About the Drug Being Reviewed

Anticipated benefits of Otezla, compared with other therapies, include the ease of oral dosing, increased treatment adherence, fewer adverse events, pain-free therapy, and a safer alternative to methotrexate and cyclosporine. Patients want more options for treating their psoriasis, as therapies may work for a limited amount of time. Otezla is seen as an additional therapy that may give patients symptom relief when other treatments cease to work.

The majority of patients who have taken Otezla saw a reduction in the scale, redness, and shedding of their lesions. Many preferred Otezla as it better managed their itch and scaling compared with other therapies, and found it was easier to maintain adherence due to the convenient oral dosing. Patients whose physical symptoms were effectively treated by Otezla felt an increased quality of life and improved self-confidence. Patients who had tried other therapies and found them to be inefficacious after a time found that having additional options, such as Otezla, is important in managing their psoriasis.

Patients are in favour of Otezla being recommended for listing on all participating drug plans.

APPENDIX 2: LITERATURE SEARCH STRATEGY

OVERVIEW	
Interface:	Ovid
Databases:	Embase 1974 to present MEDLINE Daily and MEDLINE 1946 to present MEDLINE In-Process & Other Non-Indexed Citations **Note:** Subject headings have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	December 8, 2014
Alerts:	Biweekly (twice monthly) search updates until April 8, 2015.
Study Types:	No search filters were applied
Limits:	No date or language limits were used Human filter was applied Conference abstracts were excluded

SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
.sh	At the end of a phrase, searches the phrase as a subject heading
MeSH	Medical Subject Heading
fs	Floating subheading
exp	Explode 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
.ot	Original title
.hw	Heading word; usually includes subject headings and controlled vocabulary
.pt	Publication type
.po	Population group [PsycInfo only]
.rn	CAS registry number
.nm	Name of substance word
pmez	Ovid database code; MEDLINE In-Process & Other Non-Indexed Citations, MEDLINE Daily and Ovid MEDLINE 1946 to Present
oemezd	Ovid database code; Embase 1974 to present, updated daily

MULTI-DATABASE STRATEGY	
#	**Searches**
	MEDLINE search
1	608141-41-9.rn,nm.
2	(Otezla or apremilast* or otezia or cc 10004 or cc10004).ti,ab,rn,nm,sh,hw,ot.
3	1 or 2
4	3 use pmez
	Embase search
5	*Apremilast/
6	(Otezla or apremilast* or otezia or cc 10004 or cc10004).ti,ab.
7	5 or 6
8	7 use oemezd
9	8 not conference abstract.pt.
10	4 or 9
	Remove non-human studies and duplicates
11	exp animals/
12	exp animal experimentation/ or exp animal experiment/
13	exp models animal/
14	nonhuman/
15	exp vertebrate/ or exp vertebrates/
16	animal.po.
17	or/11-16
18	exp humans/
19	exp human experimentation/ or exp human experiment/
20	human.po.
21	or/18-20
22	17 not 21
23	10 not 22
24	remove duplicates from 23

OTHER DATABASES	
PubMed	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:	November & December, 2014
Keywords:	Otezla (apremilast), psoriasis & synonyms
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" (http://www.cadth.ca/en/resources/finding-evidence-is/grey-matters), were searched:

- Health Technology Assessment Agencies
- Health Economics
- Clinical Practice Guidelines
- Drug and Device Regulatory Approvals
- Advisories and Warnings
- Drug Class Reviews
- Clinical trials
- Databases (free)
- Internet Search.

APPENDIX 3: EXCLUDED STUDIES

Reference	Reason for Exclusion
Papp K, Cather JC, Rosoph L, Sofen H, Langley RG, Matheson RT, et al. Efficacy of apremilast in the treatment of moderate to severe psoriasis: a randomised controlled trial. Lancet. 2012 Aug 25;380(9843):738-46.	**Not pivotal, not phase 3**
Strand V, Fiorentino D, Hu C, Day RM, Stevens RM, Papp KA. Improvements in patient-reported outcomes with apremilast, an oral phosphodiesterase 4 inhibitor, in the treatment of moderate to severe psoriasis: results from a phase IIb randomized, controlled study. Health Qual Life Outcomes [Internet]. 2013 [cited 2014 Dec 18];11:82. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3661377	**Not pivotal, not phase 3**
Gottlieb AB, Strober B, Krueger JG, Rohane P, Zeldis JB, Hu CC, et al. An open-label, single-arm pilot study in patients with severe plaque-type psoriasis treated with an oral anti-inflammatory agent, apremilast. Curr Med Res Opin. 2008 May;24(5):1529-38.	**Non-randomized controlled trial**
Gottlieb AB, Matheson RT, Menter A, Leonardi CL, Day RM, Hu C, et al. Efficacy, tolerability, and pharmacodynamics of apremilast in recalcitrant plaque psoriasis: a phase II open-label study. J Drugs Dermatol. 2013 Aug;12(8):888-97.	**Non-randomized controlled trial**
Papp KA, Kaufmann R, Thaci D, Hu C, Sutherland D, Rohane P. Efficacy and safety of apremilast in subjects with moderate to severe plaque psoriasis: results from a phase II, multicenter, randomized, double-blind, placebo-controlled, parallel-group, dose-comparison study. J Eur Acad Dermatol Venereol. 2013 Mar;27(3):e376-e383.	**Wrong dose**
Kavanaugh A, Mease PJ, Gomez-Reino JJ, Adebajo AO, Wollenhaupt J, Gladman DD, et al. Treatment of psoriatic arthritis in a phase 3 randomised, placebo-controlled trial with apremilast, an oral phosphodiesterase 4 inhibitor. Ann Rheum Dis [Internet]. 2014 [cited 2015 Feb 13] Jun;73(6):1020-6. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4033106	**Wrong indication**
Strand V, Schett G, Hu C, Stevens RM. Patient-reported Health-related Quality of Life with apremilast for psoriatic arthritis: a phase II, randomized, controlled study. J Rheumatol. 2013 Jul;40(7):1158-65.	**Wrong indication**

APPENDIX 4: DETAILED OUTCOME DATA

TABLE 9: OTHER EFFICACY OUTCOMES

	ESTEEM-1		ESTEEM-2	
	Apremilast N = 562	**Placebo N = 282**	**Apremilast N = 274**	**Placebo N = 137**
PASI 50				
Patients at week 16, n (%) — FAS/LOCF	330 (59)	48 (17)	152 (56)	27 (20)
Difference in proportions [95% CI][a]	41.7 [35.7 to 47.7]		35.8 [26.9 to 44.7]	
P value	*P* < 0.0001		*P* < 0.0001	
PASI 90				
Patients at week 16, n (%) [95% CI][c]	55 (9.8) [7.5 to 12.5]	1 (0.4) [0.0 to 2.0]	24 (8.8) [5.7 to 12.8]	1 (0.7) [0.0 to 4.0]
P value	NR		NR	
Pruritus VAS				
Mean (SD) at baseline	66.2 (25.5)	65.2 (24.8)	67.8 (25.2)	65.0 (26.0)
Mean (SD) change at week 16	–31.5 (32.4)	–7.3 (27.1)	–33.5 (35.5)	–12.2 (30.9)
LS mean [95% CI]	–31.5 [–34.1 to –29.0]	–7.3 [–10.9 to –3.6]	–33.5 [–37.6 to –29.4]	–12.2 [–18.0 to –6.4]
Difference in LS means [2-sided 95% CI][e]	–24.2 [–28.7 to –19.8]		–21.3 [–28.4 to –14.2]	
P value	*P* < 0.0001		*P* < 0.0001	
NAPSI score				
Overall, mean (SD) baseline	4.3 (2.0) N = 363	4.4 (2.1) N = 195	4.2 (2.1) N = 163	4.4 (2.0) N = 84
Mean % (SD) change from baseline	–22.5 (54.9)	6.5 (60.6)	–29.0 (67.5)	–7.1 (46.6)
LS mean % change from baseline [95% CI]	–22.5 [–28.6 to –16.5]	6.5 [–1.8 to 14.9]	–29.3 [–38.7 to –20.0]	–6.4 [–19.4 to 6.5]
Difference in LS means [2-sided 95% CI][i]	–29.1 [–28.6 to –16.5]		–22.9 [–38.9 to –6.9]	
2-sided *P* value	*P* < 0.0001		*P* = 0.0052	
Patients with a NAPSI 50 at week 16, n (%)	121 (33)	29 (15)	78 (45)	17 (19)
Difference in proportions [2-sided 95% CI][j]	18.5 [11.5 to 25.4]		25.9 [15.0 to 36.8]	
2-sided *P* value	*P* < 0.0001		*P* < 0.0001	
Overall number of involved nails in patients with nail psoriasis at week 16				
Mean (SD) baseline	6.5 (3.5)	7.3 (3.3)	6.3 (3.5)	6.7 (3.4)
Mean (SD) change from baseline to week 16	–0.8 (2.8)	0.1 (2.0)	–1.3 (2.9)	0.4 (2.9)
Scalp (ScPGA)				
Score 0 or 1 at week 16	174 (47) N = 374	33 (18) N = 189	72 (41) N = 176	16 (17) N = 93
Difference in proportions [2-sided 95% CI][a]	29.1 (21.7 to 36.5)		23.7 (13.1 to 34.3)	
2-sided *P* value	*P* < 0.0001		*P* < 0.0001	
Score 0, 1, or 2 at week 16	257 (69)	67 (35)	118 (67)	39 (42)

	ESTEEM-1		ESTEEM-2	
	Apremilast N = 562	**Placebo N = 282**	**Apremilast N = 274**	**Placebo N = 137**
Difference in proportions (2-sided 95% CI)	33.3 (25.0 to 41.5)		25.1 (12.9 to 37.3)	
2-sided *P* value[j]	*P < 0.0001*		*P < 0.0001*[d]	
Pulmoplantar (PPPGA)				
Score 0 or 1 at week 16, n (%)	22 (39) N = 57	8 (31) N = 26	17 (65) N = 26	5 (31) N = 16
Difference in proportions (2-sided 95% CI)	7.8 (–14.0 to 29.6)		34.1 (5.0 to 63.3)	
2-sided *P* value[j]	*P* = 0.4912		*P = 0.0315*	
Score of 0, 1, or 2	38 (67) N = 57	13 (50) N = 26	20 (77)	6 (38)
Difference in proportions (2-sided 95% CI)	16.7 (–6.1 to 39.5)		39.4 (10.7 to 68.1)	
2-sided *P* value[j]	*P* = 0.1479		*P* = 0.0106	

ANCOVA = analysis of covariance; ANOVA = analysis of variance; CI = confidence interval; FAS = full analysis set; LOCF = last observation carried forward; LS = least squares; NAPSI = Nail Psoriasis Severity Index; NR = not reported; PASI = Psoriasis Area and Severity Index; PGA = Physician Global Assessment; PPPGA = Palmoplantar Psoriasis Physician Global Assessment; ScPGA = Scalp Physician Global Assessment; SD = standard deviation; VAS = visual analogue scale.

[a] Two-sided 95% CI is based on the normal approximation. Two-sided *P* value is based on the two-sided chi-square test.

[b] Non-responder imputations were patients who discontinued prior to week 16 or whose PASI evaluations were missing.

[c] Two-sided 95% CI is based on the Clopper-Pearson method.

[d] Based on an ANCOVA model for the per cent change from baseline at week 16, with treatment group as a factor and the baseline value as a covariate. Means (LS means), difference in LS means, and *P* values were adjusted by covariate. The two-sided *P* value for slope homogeneity is > 0.05.

[e] Based on an ANOVA model for the change from baseline at week 16, with treatment group as a factor (an ANOVA model). Unadjusted means and *P* values are provided. The two-sided *P* value for slope homogeneity is < 0.05.

[f] Based on an ANCOVA model for the per cent change from baseline at week 16, with treatment group as a factor and the baseline value as a covariate. Means (LS means) and *P* values were adjusted by covariate. The two-sided *P* value for slope homogeneity is > 0.05.

[g] Nominal statistical significance due to hierarchical testing.

[h] *P* values in italics are ≤ 0.050 and considered nominally significant, as there was no adjustment for multiplicity based on hierarchical testing.

[i] Based on an ANCOVA model for the change from baseline at week 16. The model includes treatment group as a factor (ANOVA model). The unadjusted means and *P* value are provided. The two-sided *P* value for slope homogeneity is < 0.05.

[j] Two-sided 95% CI is based on the normal approximation. Two-sided *P* value is based on the two-sided chi-square test.

Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

TABLE 10: SUBGROUP DATA

	ESTEEM-1		ESTEEM-2	
	Apremilast N = 562	**Placebo N = 282**	**Apremilast N = 274**	**Placebo N = 137**
PASI 75: Patients with Response, n (%)				
History of psoriatic arthritis				
Yes	[redacted]	[redacted]	[redacted]	[redacted]
Difference in proportions (95% CI)	[redacted]		[redacted]	
No	[redacted]	[redacted]	[redacted]	[redacted]
Difference in proportions (95% CI)	[redacted]		[redacted]	
Failed prior phototherapy				
Yes	[redacted]	[redacted]	[redacted]	[redacted]
Difference in proportions (95% CI)	[redacted]		[redacted]	
No	[redacted]	[redacted]	[redacted]	[redacted]
Difference in proportions (95% CI)	[redacted]		[redacted]	
Failed conventional systemic therapy				
Yes	15/60 (25)	1/36 (3)	9/38 (24)	3/18 (17)
Difference in proportions (95% CI)	22.2 (10.0 to 34.4)		7.0 (–14.9 to 28.9)	
No	42/152 (28)	1/66 (2)	14/68 (21)	2/35 (6)
Difference in proportions (95% CI)	26.1 (18.4 to 33.8)		14.9 (2.6 to 27.2)	
Failed biologic therapy				
Yes	8/37 (22)	0/19 (0)	6/24 (25)	1/11 (9)
Difference in proportions (95% CI)	21.6 (8.4 to 34.9)		15.9 (-8.4 to 40.2)	
No	35/125 (28)	3/61 (5)	15/68 (22)	1/33 (3)
Difference in proportions (95% CI)	23.1 (13.5 to 32.6)		19.0 (7.6 to 30.5)	
Failed TNF blocker				
Yes	7/26 (27)	0/15 (0)	4/17 (24)	1/10 (10)
Difference in proportions (95% CI)	26.9 (9.9 to 44.0)		13.5 (–13.9 to 41.0)	
No	19/75 (25)	1/33 (3)	11/55 (20)	1/20 (5)
Difference in proportions (95% CI)	22.3 (10.9 to 33.8)		15.0 (0.8 to 29.2)	

PASI = Psoriasis Area and Severity Index; TNF = tumour necrosis factor.
Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

TABLE 11: EFFICACY DATA FROM MAINTENANCE PHASE

	ESTEEM-1		ESTEEM-2	
PASI 75	**APR/APR N = 562**	**PLA/APR N = 282**	**APR/APR N = 274**	**PLA/APR N = 108**
Patients with, n (%) — FAS at week 32 [95% CI]	159 (28.3) [24.6 to 32.2]	76 (31.0) [25.3 to 37.2]	68 (24.8) [19.8 to 30.4][a]	31 (28.7) [20.4 to 38.2][a]
PASI 50				
Patients with, n (%) — FAS at week 32 [95% CI]	301 (53.6) [49.3 to 57.7]	156 (63.7) [57.3 to 69.7]	126 (46.0) [40.0 to 52.1][a]	71 (65.7) [56.0 to 74.6][a]
PASI 90				
Patients with, n (%) — FAS at week 32 [95% CI]	68 (12.1) [9.5 to 15.1]	22 (9.0) [5.7 to 13.3]	26 (9.5) [6.3 to 13.6][a]	9 (8.3) [3.9 to 15.2][a]
Change in PASI				
Mean (SD) at baseline	NR	NR	NR	NR
Mean (SD) % change at week 32	–61.9 (27.8) N = 425	–62.2 (23.7) N = 216	–58.8 (28.4) N = 191	–59.8 (27.5) N = 98
sPGA				
Patients with, n (%) — FAS at week 32 [95% CI]	135 (24.0) [20.5 to 27.8]	62 (25.3) [20.0 to 31.2]	49 (17.9) [13.5 to 22.9]	25 (23.1) [15.6 to 32.2]
Affected BSA				
Mean (SD) at baseline	NR	NR	NR	NR
Mean (SD) % change at week 32	–61.2 (34.2)	–59.3 (30.8)	–60.7 (33.5) N = 191	–58.7 (38.2) N = 98
DLQI Total score				
Mean (SD) at baseline	12.7 (7.1)	11.9 (6.6)	12.6 (7.2)	12.3 (7.0)
Mean (SD) change at week 32	–7.3 (6.6) N = 422	–6.3 (5.7) N = 216	–7.0 (6.4) N = 191	–7.4 (7.2) N = 98

APR/APR = apremilast in randomized controlled trial phase followed by apremilast in maintenance phase; BSA = body surface area; CI = confidence interval; DLQI = Dermatology Life Quality Index; FAS = full analysis set; NR = not reported; PASI = Psoriasis Area and Severity Index; PLA/APR = placebo in randomized controlled trial phase followed by apremilast in maintenance phase; sPGA = static Physician Global Assessment; SD = standard deviation.

[a] Two-sided 95% CI based on the Clopper-Pearson method.

Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

TABLE 12: EFFICACY DATA FROM RANDOMIZED TREATMENT WITHDRAWAL PHASE (WEEKS 32 TO 52)

	ESTEEM-1		ESTEEM-2	
PASI 75:Time to Loss of Response, Weeks[a]	**APR/APR/APR N = 77**	**APR/APR/PLA N = 77**	**APR/APR/APR N = 36**	**APR/APR/PLA N = 31**
Patients who lost PASI 75, n (%)	40	63	17	28
Median [95% CI][b]	17.7 [13.0 to –]	5.1 [4.1 to 8.1]	20.9 [8.1 to –]	4.3 [3.0 to 8.0]
Hazard ratio [95% CI][c]	2.649 [1.768 to 3.969]		3.459 [1.835 to 6.519]	
2-sided *P* value[d, e]	*P* < 0.0001		*P* < 0.0001	
sPGA: Time to Loss of Response, Weeks [a]				
Patients who lost sPGA	24 (41.4)	41 (70.7)	11 (44.0)	19 (82.6)
Median (2-sided 95% CI)[b]	20.6 [19.1 to –]	5.0 [4.1 to 8.1]	20.9 [19.0 to 21.4]	4.0 [2.6 to 5.1]
Hazard ratio (95% CI)[c]	3.746 [2.192 to 6.400]		4.511 [1.896 to 10.729]	
2-sided *P* value[d, e]	*P* < 0.0001		*P* = 0.0002	

APR/APR/APR = apremilast in randomized controlled trial phase followed by apremilast in maintenance phase and then apremilast in randomized treatment withdrawal phase; APR/APR/PLA = apremilast in randomized controlled trial phase followed by apremilast in maintenance phase and then placebo in randomized treatment withdrawal phase; BID = twice daily; CI = confidence interval; PASI = Psoriasis Area and Severity Index; sPGA = static Physician Global Assessment.

[a] Time to loss is the time between the re-randomization date and the date of the first assessment where loss of PASI 75 is observed (event); or the time between the re-randomization date and the date of the last PASI assessment in the weeks 32 to 52 interval prior to addition of protocol-prohibited medication/therapy, or resumption of apremilast, or discontinuation, or week 52 if no loss (censored).

[b] Median (2-sided 95% CI) are based on Kaplan–Meier estimates.

[c] Hazard ratio (Placebo/APR 30 BID) (two-sided 95% CI) is based on the Cox model with treatment as the independent variable.

[d] The 2-sided *P* value for the treatment comparison is based on the log-rank test.

[e] Nominal statistical significance due to hierarchical testing.

Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

TABLE 13: HARMS DATA FROM MAINTENANCE PHASE (WEEKS 16 TO 32)

	ESTEEM-1		ESTEEM-2	
AEs	**APR/APR N = 493**	**PLA/APR N = 282**	**APR/APR N = 234**	**PLA/APR N = 108**
Patients with > 0 AEs, n (%)	274 (56)	160 (66)	129 (55)	60 (56)
Most common AEs				
Diarrhea	11 (2)	31 (13)	5 (2)	7 (7)
Nausea	13 (3)	20 (8)	6 (3)	6 (6)
URTI	47 (10)	23 (9)	6 (3)	7 (7)
SAEs				
Patients with > 0 SAEs, n (%)	8 (2)	4 (2)	5 (2)	1 (1)
WDAEs				
WDAEs, n (%)	9 (2)	10 (4)	6 (3)	2 (2)
DEATHS				
Number of deaths, n (%)	0	0	0	0
NOTABLE HARMS				
Weight loss reported as an AE	NR	NR	NR	NR
Weight loss of > 10%	[redacted]	[redacted]	[redacted]	[redacted]

AE = adverse event; APR/APR = apremilast in randomized controlled trial phase followed by apremilast in maintenance phase; NR = not reported; PLA/APR = placebo in randomized controlled trial phase followed by apremilast in maintenance phase; SAE = serious adverse events; URTI = upper respiratory tract infection; WDAE = withdrawal due to adverse event.
Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

TABLE 14: HARMS DATA FROM RANDOMIZED TREATMENT WITHDRAWAL PHASE (WEEKS 32 TO 52)

	ESTEEM-1					ESTEEM-2				
	Patients Re-randomized			Patients Not Re-randomized		Patients Re-randomized			Patients Not Re-randomized	
AEs	**APR/APR/APR N = 77**	**APR/APR/PLA N = 77**	**APR/APR/PLA/APR N = 64**	**APR/APR/APR N = 245**	**PLA/APR/APR N = 208**	**APR/APR/APR N = 61**	**APR/APR/PLA N = 62**	**APR/APR/PLA/APR N = 32**	**APR/APR/APR N = 58**	**PLA/APR/APR N = 96**
Patients with > 0 AEs, n (%)	41 (53)	27 (35)	25 (39)	127 (52)	114 (55)	35 (57)	29 (47)	17 (53)	31 (53)	51 (53)
Most common AEs[a]										
URTI	8 (10)	2 (3)	4 (6)	22 (9)	20 (10)	4 (7)	1 (2)	0	3 (5)	6 (6)
Nasopharyngitis	6 (8)	3 (4)	1 (2)	18 (7)	7 (4)	4 (7)	5 (8)	2 (6)	4 (7)	7 (7)
SAEs										
Patients with > 0 SAEs, n (%)	0	2 (3)	1 (2)	5 (2)	2 (1)	1 (2)	2 (3)	0	2 (3)	3 (3)
WDAEs										
WDAEs, n (%)	1 (1)	0	0	4 (2)	6 (3)	1 (2)	2 (3)	0	0	1 (1)
Deaths										
Number of deaths, n (%)	0	0	0	0	0	0	1	0	0	0
Notable Harms										
Weight loss reported as an AE	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Weight loss > 10%	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhea	2 (3)	0	2 (3)	6 (2)	8 (4)	2 (3)	1 (2)	1 (3)	1 (2)	2 (2)
Nausea	2 (3)	0	1 (2)	6 (2)	3 (1)	3 (5)	2 (3)	2 (6)	1 (2)	2 (2)

AE = adverse event; APR/APR/APR = apremilast in all phases; APR/APR/PLA = apremilast in randomized controlled trial phase followed by apremilast in maintenance phase and then placebo in randomized treatment withdrawal phase; APR/APR/PLA/APR = apremilast in randomized controlled trial phase followed by apremilast in maintenance phase and then placebo in randomized treatment withdrawal phase, who then had to resume apremilast; CI = confidence interval; PASI = Psoriasis Area and Severity Index; PLA/APR/APR = placebo in randomized controlled trial phase followed by apremilast in maintenance phase and then placebo in randomized treatment withdrawal phase; SAE = serious adverse event; sPGA = static Physician Global Assessment; URTI = upper respiratory tract infection; WDAE = withdrawal due to adverse events.
Source: Clinical Study Reports for ESTEEM-1 and ESTEEM-2.[3,4]

APPENDIX 5: VALIDITY OF OUTCOME MEASURES

Aim

To summarize the validity of the following outcome measures:

- Body surface area (BSA)
- Dermatology Life Quality Index (DLQI)
- EuroQol 5-Dimensions Health-Related Quality of Life Questionnaire (EQ-5D)
- Psoriasis Area and Severity Index (PASI)
- Physician Global Assessment (PGA).

Findings

TABLE 15: VALIDITY AND MINIMAL CLINICALLY IMPORTANT DIFFERENCE OF OUTCOME MEASURES

Instrument	Type	Evidence of Validity	MCID	References
BSA	BSA is a clinician-measured estimate of the percentage of a patient's body surface affected by psoriasis.	Yes	Unknown	Hani et al. 2012,[12] Schmitt and Wozel 2005[13]
DLQI	DLQI is a 10-item, dermatology-specific quality of life questionnaire.	Yes	3.2	Mattei et al. 2014,[14] Ruderman et al. 2003,[15] Shikiar et al. 2006[16]
EQ-5D	EQ-5D is a general, non-disease-specific health-related quality of life questionnaire.	Yes	0.09 to 0.20	Norlin et al. 2012,[17] Shikiar et al. 2006[16]
PASI	Numeric score ranging from 0 to 72, based on assessments of 4 body areas and severity of induration, erythema, and scaling.	Yes	Unknown	Ashcroft et al. 1999,[18] Carlin et al. 2004,[19] Feldman et al. 2004,[20] Gourraud et al. 2012[21]
PGA	Single estimate of a patient's disease severity at a given time based on induration, erythema, and scaling.	Yes	Unknown	Feldman et al. 2004,[20] Weisman et al. 2003[22]

BSA = body surface area; DLQI = Dermatology Life Quality Index; EQ-5D = EuroQol 5-Dimensions Health-Related Quality of Life Questionnaire; MCID = minimal clinically important difference; PASI = Psoriasis Area and Severity Index; PGA = Physician Global Assessment.

Body surface area

Measured by clinicians, BSA provides an estimate of the percentage of a patient's body surface affected by psoriasis; this measure ranges from 0% to 100%.[23] To assist in BSA estimation, the patient's palm is estimated to be 1% of the total BSA.[24,25] BSA assessment has been validated on medical mannequins[12] and is used for the calculation of PASI scores.[24,26] However, several limitations exist, including high inter-rater variability; it does not consider lesion severity; and often BSA is overestimated.[13] As a single instrument, BSA is thought to be insufficient in measuring psoriasis severity, as the measure does not take into account the severity of lesions.[13]

Dermatology Life Quality Index

The DLQI is a widely used dermatology-specific quality of life instrument. It is a 10-item questionnaire that assesses six different aspects that may affect quality of life.[27] These aspects are symptoms and feelings, daily activities, leisure, work and school performance, personal relationships and treatment. The maximum score per aspect is either 3 or 6, and the scores for each can be expressed as a percentage of either 3 or 6. Each of the 10 questions is scored from 0 (not at all) to 3 (very much), and the overall DLQI is calculated by summing the score of each question resulting in a numeric score between 0 and 30 (or a percentage of 30). [27] The higher the score, the more quality of life is impaired. The meaning of the DLQI scores on a patient's life is as follows: [28]

- 0 to 1 = no effect
- 2 to 5 = small effect
- 6 to 10 = moderate effect
- 11 to 20 = very large effect
- 21 to 30 = extremely large effect

The DLQI has shown good reliability and construct validity.[15] The estimated minimal clinically important difference (MCID) for the DLQI in patients with psoriasis is 3.2.[14] Estimates of the minimal important difference (the smallest difference a patient would regard as beneficial) have ranged from 2.3 to 5.7.[16]

- A number of limitations of the DLQI have been identified: Concerns have been identified regarding unidimensionality and the behaviour of items of the DLQI in different psoriatic patient populations with respect to their age, gender, culture, etc.[28]
- The patient's emotional aspects may be underrepresented, and this may be one reason for unexpectedly low DLQI scores in patients with more emotionally disabling diseases such as vitiligo. To overcome this, it is suggested that the DLQI be combined with more emotionally oriented measures such as the mental component of the Short-Form 36-item Health Survey (SF-36) or Hospital Anxiety and Depression (HAD) scales.[28]
- The non-availability of benchmarks for the MCID of DLQI scores in general dermatological conditions is a limitation that has been identified, although there have been some attempts to determine these differences for specific conditions such as psoriasis.[28]
- DLQI may lack sensitivity in detecting change from mild to severe psoriasis.[29]

EuroQol 5-Dimensions Health-Related Quality of Life Questionnaire

The EQ-5D is a generic quality of life instrument that may be applied to a wide range of health conditions and treatments.[30] The first of two parts of the EQ-5D is a descriptive system that classifies respondents (aged 12 years or older) into one of 243 distinct health states. The descriptive system consists of the following five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has three possible levels (1, 2, or 3) representing "no problems," "some problems," and "extreme problems," respectively. Respondents are asked to choose the level that reflects their health state for each of the five dimensions. A scoring function can be used to assign a value (EQ-5D index score) to self-reported health states from a set of population-based preference weights.[30] The second part is a 20 cm visual analogue scale (EQ-VAS) that has end points labelled 0 and 100, with respective anchors of "worst imaginable health state" and "best imaginable health state." Respondents are asked to rate their health by drawing a line from an anchor box to the point on the EQ-VAS that best represents their health on that day. Hence, the EQ-5D produces three types of data for each respondent:

1. A profile indicating the extent of problems on each of the five dimensions represented by a five-digit descriptor, such as 11121, 33211, etc.

2. A population preference-weighted health index score based on the descriptive system
3. A self-reported assessment of health status based on the EQ-VAS.

The EQ-5D index score is generated by applying a multi-attribute utility function to the descriptive system. Different utility functions are available that reflect the preferences of specific populations (e.g., US or UK). The lowest possible overall score (corresponding to severe problems on all five attributes) varies depending on the utility function that is applied to the descriptive system (e.g., –0.59 for the UK algorithm and –0.109 for the US algorithm). Scores less than 0 represent health states that are valued by society as being worse than death, while scores of 0 and 1.00 are assigned to the health states "dead" and "perfect health," respectively. Reported clinically important differences (CIDs) for this scale have ranged from 0.033 to 0.074.[31] The CIDs were derived from patients with a variety of chronic and acute conditions, including rheumatoid arthritis, osteoarthritis, irritable bowel syndrome, and acute myocardial infarction.[32,33]

The evidence for the validity of EQ-5D in the psoriasis population is limited. A Swedish observational cohort study found good correlation of EQ-5D with other outcome measures DLQI and PASI.[17] However, EQ-5D was not as responsive to change in patients' clinical status as the DLQI, and the study authors recommend the use of EQ-5D to complement DLQI and PASI.[17] An additional study found the EQ-5D to be highly correlated with the DLQI, although not as responsive to change in patient status.[16] EQ-5D showed similar responsiveness as another quality of life measure: the SF-36.[16] Estimates of the MCID for EQ-5D were derived using distributional and anchor-based approaches.[16] PASI and PGA were used in the anchor-based approach for determining MCID, and are as follows: an estimate obtained using PASI scores of near-responders (PASI 25 to PASI 49); an estimated obtained using the PASI scores of partial responders (PASI 50 to PASI 74); and an estimate based on the difference between non-responders and minimal responders using the PGA.[16] The estimated MCID for EQ-5D in the psoriasis population ranges from 0.09 to 0.20.[16] This estimated MCID, compared with the general MCID range of 0.033 to 0.074, suggests that a larger difference in EQ-5D score is necessary for psoriasis patients to regard the change as clinically beneficial.

Physician's Global Assessment

The PGA is used to determine a single estimate of the patient's overall severity of disease at a given point in time. Various PGAs have been used in psoriasis with different descriptions and scores.[25] The static PGA (sPGA) was used in the current submission for Otezla; this evaluates PGA at the time of patient assessment.[4] The sPGA used for this review was a five-point scale (0 = clear, to 4 = severe), based on scores for the assessed severity of erythema, plaque elevation (induration), and scaling.[4]

Psoriatic lesions are graded for induration, erythema, and scaling based on scales of 0 to 4 that are then averaged over all lesions.[26] For induration (I), the scale denotes the extent of plaque elevation as follows: 0 = no evidence, 1 = minimal, 2 = mild or slight, 3 = elevated, 4 = marked.[26] For erythema (E), 0 = no evidence of erythema although hyperpigmentation may be present, 1 = faint erythema, 2 = light red colouration, 3 = red colouration, 4 = dark to deep red colouration.[26] For scaling (S), 0 = no evidence of scaling, 1 = minimal; occasional fine scale, 2 = fine scale dominates; 3 = moderate; coarse scale predominates; 4 = marked; thick, non-tenacious scale dominates.[26]

The sum of the three scales are added and then divided by three (I + E + S/3) to obtain a final PGA score as follows:

- 0 = cleared, except for residual discolouration
- 1 = minimal — majority of lesions have individual scores for I + E + S/3 that averages 1

- 2 = mild — majority of lesions have individual scores that averages 2
- 3 = moderate — majority of lesions have individual scores that averages 3
- 4 = severe — majority of lesions have individual scores that averages 4

The PGA is more subjective than PASI in that there is no attempt to quantify the individual elements of plaque morphology or BSA involvement.[20,22] There have also been fewer studies using PGA than PASI. This outcome is considered reliable using test–retest data and internal consistency.[22] However, inter-rater reliability due to variability, especially in untrained observers, is poor.[22] Many studies now employ only the final value of "clear" or "almost clear" as treatment success. Although it would seem that the PGA may be less likely to be open to interpretation, different studies have used different definitions of "clear" or "almost clear," making comparisons between treatments difficult.[22] Construct and content validity are considered strong within a study, but comparison with other studies as well as relationship to other methods are problematic due to the variability in data collection, analysis, and reporting method.[22]

Psoriasis Area and Severity Index

PASI is a widely used instrument in psoriasis trials that assesses and grades the severity of psoriatic lesions and the patient's response to treatment. It produces a numeric score ranging from 0 to 72. In general, a PASI score of 5 to 10 is considered moderate disease, and a score of more than 10 is considered severe. A 75% reduction in the PASI score (PASI 75) is the current benchmark for most clinical trials in psoriasis and the criterion for efficacy of new psoriasis treatments approved by the FDA.[19]

In calculating the PASI score, severity is determined by dividing the body into four regions: head (h), upper extremities (u), trunk (t), and lower extremities (l) that account for 10%, 20%, 30%, and 40% of the total BSA, respectively.[27] Each of these areas is assessed separately for erythema, induration, and scaling, which is rated on a scale of 0 (none) to 4 (very severe). Extent of psoriatic involvement is graded as follows: 0 = no involvement; 1 = 1% to 9%; 2 = 10% to 29%; 3 = 30% to 49%; 4 = 50% to 69%; 5 = 70% to 89%; and 6 = 90% to 100%. The following formula is used to calculate the PASI score:

$$PASI = 0.1 (E_h + I_h + S_h) A_h + 0.2 (E_u + I_u + S_u) A_u + 0.3 (E_t + I_t + S_t) A_t + 0.4 (E_l + I_l + S_l) A_l$$ [27]

Where E = erythema, I = induration, S = scaling, A = area, h = head score, t = trunk score, u = upper extremities, and l = lower extremities score. PASI 75 is a dichotomous scale (Yes/No, patient achieved ≥ 75% improvement from baseline PASI score).

A number of limitations of PASI have been identified:

- PASI has been criticized as not correlating the clinical extent of the disease with quality of life and the psychological stress caused by psoriasis. The patient's measure of quality of life is often worse than the physician's-rated clinical severity.[34]
- There are significant inter-rater reliability issues regarding the measurement of BSA.[18,20] There has been some work regarding the development of imaging and analysis systems to objectively measure BSA.[12]
- PASI scores can vary substantially between experienced and inexperienced physicians, raising concerns for inter-rater reliability.[25]

- Improvements in PASI score are not linearly related to severity or improvements in psoriasis.[19,20] The extent of psoriatic involvement is measured using a scale of 1 to 6, and the areas corresponding to each score are nonlinear.
- Some severe disease (clinically) may be scored low. For example, scores as low as 3 (on palms and soles) may represent psoriasis that disables a patient from work and other life activities.
- Most patients fall into a narrow band of scores, thereby decreasing the usefulness of the full range of scores (i.e., scores above 40 are rare).[18] Validity of this scale may be overrated, in part because of the skew toward lower scores.[21]
- There is little research on the reliability of the assessments for erythema, desquamation, and induration, together with overall PASI scores.[18]
- Criterion validity is restricted by the lack of a "gold standard" measure of psoriatic severity.[35]
- PASI lacks sensitivity as erythema, desquamation, and induration are scored with equal weight within each of the four body regions. Thus, a reduction in scaling with a concomitant increase in skin erythema could be recorded with the same PASI score.
- Improvement of the histological phenotype of psoriasis can be underestimated by the per cent improvement in PASI (e.g., reduction of T cells, loss of K16 expression, and reduction in epidermal thickness).[19]
- Little work has been done to determine the clinical relevance of derived PASI scores.[18]

Conclusion

Several instruments are used when assessing psoriasis disease severity. PASI, of which BSA assessment is an integral part, is one of the most widely used tools. While there are some noted limitations of PASI, it is considered the gold standard for measuring severity of psoriasis.[12,13,36] The PGA is also widely used, although it is thought to be more subjective than PASI and does not include an assessment of affected BSA. Various PGA scales exist, and while the validity of the scale is good within a study, comparison to other studies is not always possible.

Quality of life measures are also important in the assessment of psoriasis severity. The DLQI is a dermatology-specific quality of life measure. DLQI has been validated for use in the psoriasis patient population, with an estimated MCID of 3.2.[14] The EQ-5D is a general-health specific quality of life measure. There is evidence for the concurrent validity of the EQ-5D in the psoriasis patient population, as it correlates well with DLQI and PASI. Quality of life remains an important consideration for assessing severity of disease for patients with psoriasis.

APPENDIX 6: SUMMARY OF OTHER STUDIES

Objective

To summarize the results of the ESTEEM-1 open-label extension study.

Findings

Main findings from ESTEEM-1, excluding the extension phase, are presented in the main body of this report. Data from the manufacturer-provided ESTEEM-1 open-label trial[37] are summarized below. The main purpose of the provided data was to report the safety findings from the 104-week period. Results are presented from weeks 0 to ≤ 52 (804 patients randomized), and from the extension study weeks > 52 to ≤ 104 (444 patients).

TABLE 16: SUMMARY OF SAFETY DATA FROM WEEKS 0 TO ≤ 104

	Weeks 0 to ≤ 52	Weeks > 52 to ≤ 104
AEs, n (%)	**APR 30 mg b.i.d. (N = 804; 624.4 patient-years)**	**APR 30 mg b.i.d. (N = 444; 350.7 patient-years)**
Diarrhea	150 (18.7)	8 (1.8)
Nausea	122 (15.2)	3 (0.7)
URTI	146 (18.2)	43 (9.7)
Nasopharyngitis	110 (13.7)	30 (6.8)
Tension headache	77 (9.6)	7 (1.6)
Headache	52 (6.5)	3 (0.7)
WDAEs, n (%)	63 (7.8)	13 (2.9)
Severe AEs, n (%)	50 (6.2)	14 (3.2)
SAEs, n (%)	36 (4.6)	24 (5.4)
Deaths, N	1	1

AE = adverse event; APR = apremilast; b.i.d. = twice daily; SAE = serious adverse event; URTI = upper respiratory tract infection; WDAE = withdrawal due to adverse event.

The cause of death for the patient who died during weeks > 52 to ≤ 104 was a cerebrovascular accident; this patient had a prior history of diabetes mellitus, hypertension, and hyperlipidemia. The cause of death was reported as cardiac failure for the patient who died during weeks 0 to ≤ 52.

Other notable harms are as follows: during weeks > 52 to ≤ 104, six patients experienced serious infection, although none were opportunistic. The authors of the study also noted that no apremilast-treated patients experienced reactivation of tuberculosis. Although 2.0% of patients experienced depression during weeks 0 to ≤ 52, and 0.5% experienced depression during weeks > 52 to ≤ 104, there was no reported suicidal ideation, or attempted suicide. Baseline weight was available for 473 patients during weeks 0 to ≤ 52 (mean baseline = 93.54 kg), and for 202 patients for weeks > 52 to ≤ 104. For weeks 0 to ≤ 52, mean and median weight change from baseline was –2.07% and –1.59%, respectively. For weeks > 52 to ≤ 104, mean and median weight change from baseline were –1.94% and –1.56%, respectively.

Summary

The trial was primarily designed to investigate safety of apremilast. Main findings are summarized for weeks 0 to ≤ 52, and the extension period of weeks > 52 to ≤ 104. Design constraints for the extension period of ESTEEM-1 (e.g., open-label, no control group) limit the usefulness for providing any further information on the efficacy for apremilast.

APPENDIX 7: SUMMARY OF COMPARATORS

Introduction

Neither of the pivotal phase 3 studies, ESTEEM-1 and ESTEEM-2, included in the manufacturer's submission were designed to compare apremilast to active comparators. Therefore, the manufacturer submitted a network meta-analysis (NMA) in order to obtain comparative efficacy data.

Objectives

To provide a summary and critical appraisal of the manufacturer-submitted indirect treatment comparison and NMA.

Methods

Eligibility Criteria

[redacted]

Review Methods

[redacted]

Population

[redacted]

Interventions and Comparators

[redacted]

TABLE 17: SUMMARY OF DOSING IN NETWORK META-ANALYSIS

[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

Outcomes

[redacted]

Critical Appraisal

[redacted]

Evidence Network

FIGURE 2: EVIDENCE NETWORK

Figure 3 contained confidential information and was removed at the request of the manufacturer.

Note: Each node represents a treatment and/or a dosing regimen, and each link connects treatments that have been directly compared in one or more randomized clinical trials. Thicker lines represent higher number of comparisons in the network.

BID = twice daily; BIW = twice weekly.
Source: CADTH Common Drug Review submission.[9]

Analysis

The following is an excerpt from the NMA submitted by the manufacturer:[38]

[redacted]

[39,40]

[redacted]

Results

Study and Patient Characteristics

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Results of the Meta-analysis

[redacted]

TABLE 18: SUMMARY OF PSORIASIS AREA AND SEVERITY INDEX DATA: ENTIRE POPULATION

Probability of:	PASI 75	PASI 90	PASI 50
[redacted]			

BID = twice daily; BIW = twice weekly; CrI = credible interval; EOW = every other week; PASI = Psoriasis Area and Severity Index; Q12W = every 12 weeks; SD = standard deviation; wks = weeks.

TABLE 19: SUMMARY OF PSORIASIS AREA AND SEVERITY INDEX DATA: BIOLOGIC-NAIVE SUBGROUP

Probability of:	PASI 75	PASI 90	PASI 50
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

BID = twice daily; BIW = twice weekly; CrI = credible interval; EOW = every other week; PASI = Psoriasis Area and Severity Index; Q12W = every 12 weeks; SD = standard deviation; wks = weeks.

FIGURE 3: ODD RATIOS FOR PASI 75 FOR ALL TREATMENTS COMPARED WITH PLACEBO

Figure 4 contained confidential information and was removed at the request of the manufacturer.

NMA = network meta-analysis; PASI = Psoriasis Area and Severity Index; vs. = versus.
Source: CADTH Common Drug Review submission.[9]

FIGURE 4: ODD RATIOS FOR PASI 90 FOR ALL TREATMENTS COMPARED WITH PLACEBO

Figure 5 contained confidential information and was removed at the request of the manufacturer.

NMA = network meta-analysis; PASI = Psoriasis Area and Severity Index; vs. = versus.
Source: CADTH Common Drug Review submission.[9]

FIGURE 5: ODD RATIOS FOR PASI 50 FOR ALL TREATMENTS COMPARED WITH PLACEBO

Figure 6 contained confidential information and was removed at the request of the manufacturer.

NMA = network meta-analysis; PASI = Psoriasis Area and Severity Index; vs. = versus.
Source: CADTH Common Drug Review submission.[9]

TABLE 20: SUMMARY OF BASELINE CHARACTERISTICS FROM INCLUDED STUDIES

Table 20 contained confidential information and was removed at the request of the manufacturer.

BSA = body surface area; BID = twice daily; BIW = twice weekly; EOW = every other week; NS = not stated; PASI = Psoriasis Area and Severity Index; PGA = Physician Global Assessment; PSO = psoriasis; PSSI = Psoriasis Scalp Severity Index; SD = standard deviation.
[a] Data presented as mean ± standard error (SE).

Critical Appraisal of Network Meta-analysis

Strengths

The study rationale and objectives were clearly stated in the introductory sections.

The authors stated the eligibility criteria for inclusion into the NMA, as well as their search strategy and methods for selecting studies. The data extraction strategy was also described. Treatment groups were selected that employed the recommended or usual dose of a given drug. The NMA focused on PASI scores (PASI 50/75/90) for assessment of effectiveness, and the manufacturer rationale for using PASI was that it was the primary clinical response criteria used in the economic model. PASI is considered the standard for assessing efficacy in clinical trials for psoriasis, although this was not highlighted by the authors. The authors employed a Bayesian NMA, an appropriate and well-reported method, and included a figure that summarized the network. Heterogeneity was assessed and sensitivity analyses were performed (removal of outlying trials) to confirm the source of the heterogeneity, and possible

explanations for the discrepant findings from these trials were provided. Both fixed- and random-effects models were evaluated and deviance information criteria were employed to identify choice of model.

Limitations

An important limitation of the included NMA was the lack of safety data. Harms were not identified as an outcome of interest by the authors; however, safety and tolerability have been key issues over the history of drugs used for psoriasis, whether small molecules such as methotrexate or cyclosporine, or the biologics. The manufacturer stated that these were not assessed because they were not included in the economic model. There were differences between studies in the time points used to assess PASI, and this appears to be the reason why the authors reported PASI in the range of 10 to 16 weeks. This appears to be a fairly wide range, and does lead to some uncertainty in the analysis and potential heterogeneity, and in generalizing results to the apremilast studies included in the CADTH Common Drug Review (CDR) review, which assessed PASI at 16 weeks. The authors did employ scatter plots to identify outlying data with respect to outcome data over time, and concluded that outcomes measured during weeks 10 to 16 could be compared directly.

The analysis presented by the manufacturer did not provide estimates of effects for the comparison of apremilast to other drugs, and instead compared each drug to placebo. These findings were only presented in a graph. Data for all drugs were presented in rank order, and therefore the analysis could not go beyond simply stating that one therapy ranked higher than another.

Although there were 29 trials included in this NMA, these include eight different drugs, and therefore there were a relatively low number of included trials per drug, perhaps introducing uncertainty into the analysis. Of the 29 included studies, only a slight majority (58%) were rated as either of “excellent” or “good” quality, and 28% were rated as “poor” quality. The number of studies with a rating of poor decreased to 16% when the requirement for reporting power calculations was relaxed.

Summary

Due to a lack of direct comparisons between drugs for psoriasis, the manufacturer submitted an indirect comparison, a NMA that included apremilast, other small molecules such as methotrexate and cyclosporine, as well as biologics such as infliximab, adalimumab, ustekinumab, and etanercept. The NMA included a wide variety of key comparators for apremilast, and assessed PASI responses, a key outcome in the CDR review. The major limitation, however, was that the NMA did not compare apremilast directly to other drugs, and instead compared all drugs to placebo. Therefore, the NMA does not provide any insight into how apremilast compares to other drugs for this indication. Other limitations included the lack of assessment of safety, and other efficacy outcomes.

TABLE 21: CRITICAL APPRAISAL BASED ON ISPOR NETWORK META-ANALYSIS CHECKLIST

Checklist Item	Details and Comments
Are the rationale for the study and the study objectives stated clearly?	Rationale clearly stated
Does the methods section include the following? Description of eligibility criteria Information sources Search strategy Study selection process	Search strategy and databases presented Inclusion criteria presented Number of trials meeting eligibility criteria were clearly presented Data extraction methods explained Assessment of bias was performed, and details were

Checklist Item	Details and Comments
Data extraction (validity/quality assessment of individual studies)	provided (checklist)
Are the outcome measures described?	Efficacy outcomes of interest were clearly described and rationale for their selection was explained Harms were not included as an outcome, only PASI
Is there a description of methods for analysis/synthesis of evidence? Do the methods described include the following? Description of analyses methods/models Handling of potential bias/inconsistency Analysis framework	Descriptions of analysis methods and models provided There was no clear description regarding the handling of bias
Are sensitivity analyses presented?	There were some sensitivity analyses; however, the only detailed data presented were for a subgroup analysis
Do the results include a summary of the studies included in the network of evidence? Individual study data? Network of studies?	Individual study data were summarized Network diagram was provided Forest plots are provided for each primary outcome of interest (but did not include numbers for results)
Does the study describe an assessment of model fit? Are competing models being compared?	• Authors described using model fit statistics to evaluate fixed versus random effects models
Are the results of the evidence synthesis (ITC/MTC) presented clearly?	• The only comparative data presented were each active drug versus placebo, and these were presented only as an odds ratio in a graph, without numbers. The only numbers presented were probabilities of each of PASI 50/75/90 for each drug. Therefore, no comparisons between active drugs were made.
Does the discussion include the following? • Description/summary of main findings • Internal validity of analysis • External validity • Implications of results for target audience	• The discussion was brief (< 1 page) and focused largely on strengths and limitations. • The authors appear to draw conclusions about the relative effectiveness of active drugs without ever having made these comparisons in their analyses. • The authors also appear to be more focused on comparing biologics to oral systemic therapies, versus focusing on comparing apremilast to other therapies.

ISPOR = International Society for Pharmacoeconomics and Outcomes Research; ITC = indirect treatment comparison; MTC = mixed treatment comparison; PASI = Psoriasis Area and Severity Index.

REFERENCES

1. Canadian Dermatology Association [Internet]. Ottawa (ON): Canadian Dermatology Association; c2015. Psoriasis; 2015 [cited 2015 Jan 27]. Available from: http://www.dermatology.ca/skin-hair-nails/skin/psoriasis/#!/skin-hair-nails/skin/psoriasis/living-with-psoriasis/
2. Levy AR, Davie AM, Brazier NC, Jivraj F, Albrecht LE, Gratton D, et al. Economic burden of moderate to severe plaque psoriasis in Canada. Int J Dermatol. 2012 Dec;51(12):1432-40.
3. Clinical study report no. CC-10004-PSOR-009. A phase 3, multicenter, randomized, double-blind, placebo-controlled, efficacy and safety study of apremilast (CC-10004) in subjects with moderate to severe plaque psoriasis [**CONFIDENTIAL** internal manufacturer's report]. Summit (NJ): Celgene Corporation; 2013 Jul 8.
4. Clinical study report no. CC-10004-PSOR-008. A phase 3, multicenter, randomized, double-blind, placebo-controlled, efficacy and safety study of apremilast (CC-10004) in subjects with moderate to severe plaque psoriasis [**CONFIDENTIAL** internal manufacturer's report]. Summit (NJ): Celgene Corporation; 2013 Jul 3.
5. Apremilast: patient drug information. 2014 [cited 2015 Jan 30]. In: UpToDate [Internet]. Version 15.0. Waltham (MA):1992 - . Available from: www.uptodate.com Subscription required.
6. Feldman SR. Treatment for psoriasis. 2014 [cited 2015 Jan 30]. In: UpToDate [Internet]. Version 15.0. Waltham (MA):1992 - . Available from: www.uptodate.com Subscription required.
7. e-CPS [Internet]. Ottawa: Canadian Pharmacists Association; 2015 [cited 2015 Jan 30]. Available from: https://www.e-therapeutics.ca Subscription required.
8. Health Canada reviewer's report: Otezla (apremilast) tablets, 10 mg, 20 mg and 30 mg [**CONFIDENTIAL** internal report]. Ottawa: Therapeutics Products Directorate, Health Canada; 2014.
9. CDR submission: Otezla® (apremilast), 10mg, 20mg, 30 mg tablets. Company: Celgene Inc. [**CONFIDENTIAL** manufacturer's submission]. Mississauga (ON): Celgene Inc.; 2014 Oct.
10. Rich P, Scher RK. Nail Psoriasis Severity Index: a useful tool for evaluation of nail psoriasis. J Am Acad Dermatol. 2003 Aug;49(2):206-12.
11. Manufacturer comments on CDR clinical and pharmacoeconomic reviewers' reports apremilast (Otezla®) [**CONFIDENTIAL** manufacturer's report]. Mississauga (ON): Celgene Inc.; 2015 Feb 26.
12. Hani AF, Prakasa E, Nugroho H, Affandi AM, Hussein SH. Body surface area measurement and soft clustering for PASI area assessment. Conf Proc IEEE Eng Med Biol Soc. 2012;4398-401.
13. Schmitt J, Wozel G. The psoriasis area and severity index is the adequate criterion to define severity in chronic plaque-type psoriasis. Dermatology. 2005;210(3):194-9.
14. Mattei PL, Corey KC, Kimball AB. Psoriasis Area Severity Index (PASI) and the Dermatology Life Quality Index (DLQI): the correlation between disease severity and psychological burden in patients treated with biological therapies. J Eur Acad Dermatol Venereol. 2014 Mar;28(3):333-7.
15. Ruderman EM, Markenson JA. Granulomatous infections and tumor necrosis factor antagonist therapies [abstract]. Poster presented at: EULAR; 2003 Jun 18; Lisbon, Portugal.
16. Shikiar R, Willian MK, Okun MM, Thompson CS, Revicki DA. The validity and responsiveness of three quality of life measures in the assessment of psoriasis patients: results of a phase II study. Health Qual Life Outcomes [Internet]. 2006 [cited 2015 Jan 12];4:71. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1615869/pdf/1477-7525-4-71.pdf

17. Norlin JM, Steen CK, Persson U, Schmitt-Egenolf M. Analysis of three outcome measures in moderate to severe psoriasis: a registry-based study of 2450 patients. Br J Dermatol. 2012 Apr;166(4):797-802.

18. Ashcroft DM, Wan Po AL, Williams HC, Griffiths CE. Clinical measures of disease severity and outcome in psoriasis: a critical appraisal of their quality. Br J Dermatol. 1999 Aug;141(2):185-91.

19. Carlin CS, Feldman SR, Krueger JG, Menter A, Krueger GG. A 50% reduction in the Psoriasis Area and Severity Index (PASI 50) is a clinically significant endpoint in the assessment of psoriasis. J Am Acad Dermatol. 2004 Jun;50(6):859-66.

20. Feldman SR, Menter A, Koo JY. Improved health-related quality of life following a randomized controlled trial of alefacept treatment in patients with chronic plaque psoriasis. Br J Dermatol. 2004 Feb;150(2):317-26.

21. Gourraud PA, Le GC, Puzenat E, Aubin F, Ortonne JP, Paul CF. Why statistics matter: limited inter-rater agreement prevents using the psoriasis area and severity index as a unique determinant of therapeutic decision in psoriasis. J Invest Dermatol. 2012 Sep;132(9):2171-5.

22. Weisman S, Pollack CR, Gottschalk RW. Psoriasis disease severity measures: comparing efficacy of treatments for severe psoriasis. J Dermatolog Treat. 2003 Sep;14(3):158-65.

23. Naldi L. Scoring and monitoring the severity of psoriasis. What is the preferred method? What is the ideal method? Is PASI passe? facts and controversies. Clin Dermatol. 2010 Jan;28(1):67-72.

24. Berth-Jones J, Grotzinger K, Rainville C, Pham B, Huang J, Daly S, et al. A study examining inter- and intrarater reliability of three scales for measuring severity of psoriasis: Psoriasis Area and Severity Index, Physician's Global Assessment and Lattice System Physician's Global Assessment. Br J Dermatol. 2006 Oct;155(4):707-13.

25. Langley RG, Ellis CN. Evaluating psoriasis with Psoriasis Area and Severity Index, Psoriasis Global Assessment, and Lattice System Physician's Global Assessment. J Am Acad Dermatol. 2004 Oct;51(4):563-9.

26. Cappelleri JC, Bushmakin AG, Harness J, Mamolo C. Psychometric validation of the physician global assessment scale for assessing severity of psoriasis disease activity. Qual Life Res. 2013 Nov;22(9):2489-99.

27. Feldman SR, Krueger GG. Psoriasis assessment tools in clinical trials. Ann Rheum Dis [Internet]. 2005 [cited 2014 May 21];64(Suppl 2):ii65-ii68. Available from: http://ard.bmj.com/content/64/suppl_2/ii65.full.pdf

28. Basra MKA, 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.

29. Fernandez-Penas P, Jones-Caballero M, Espallardo O, Garcia-Diez A. Comparison of Skindex-29, Dermatology Life Quality Index, Psoriasis Disability Index and Medical Outcome Study Short Form 36 in patients with mild to severe psoriasis. Br J Dermatol. 2012 Apr;166(4):884-7.

30. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy. 1990 Dec;16(3):199-208.

31. Melzer N, Meuth SG. Disease-modifying therapy in multiple sclerosis and chronic inflammatory demyelinating polyradiculoneuropathy: common and divergent current and future strategies. Clin Exp Immunol. 2014 Mar;175(3):359-72.

32. Sullivan PW, Lawrence WF, Ghushchyan V. A national catalog of preference-based scores for chronic conditions in the United States. Med Care. 2005 Jul;43(7):736-49.

33. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life Res. 2005 Aug;14(6):1523-32.

34. Choi J, Koo JY. Quality of life issues in psoriasis. J Am Acad Dermatol. 2003 Aug;49(2 Suppl):S57-S61.

35. Jacobson CC, Kimball AB. Rethinking the Psoriasis Area and Severity Index: the impact of area should be increased. Br J Dermatol. 2004 Aug;151(2):381-7.

36. Robinson A, Kardos M, Kimball AB. Physician Global Assessment (PGA) and Psoriasis Area and Severity Index (PASI): why do both? A systematic analysis of randomized controlled trials of biologic agents for moderate to severe plaque psoriasis. J Am Acad Dermatol. 2012 Mar;66(3):369-75.

37. Celgene response to January 13, 2015 CDR request for additional information regarding the Otezla CDR review: details of data from long-term extension phase of ESTEEM 1 and ESTEEM 2 trials [**CONFIDENTIAL** additional manufacturer's information]. Mississauga (ON): Celgene Inc.; 2015 Jan 26.

38. PSO Network Meta-Analysis Report - a supplemental document to the Otezla® (apremilast) pharmacoeconomic evaluation report. In: CDR submission: Otezla® (apremilast), 10mg, 20mg, 30 mg tablets. Company: Celgene Inc. [**CONFIDENTIAL** manufacturer's submission]. Mississauga (ON): Celgene Inc.; 2014 Oct.

39. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2004 Oct 30;23(20):3105-24.

40. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput. 2000;10(4):325-37.