Supplemental Appendix

Online Resource 1. PubMed Literature Search Strategy

Search Number	Search Terms	Number of Results
Disease/Condition		
#1	"Dermatitis, Atopic"[Mesh] OR "atopic dermatitis"[Text Word] OR "atopic eczema"[Text Word] OR "hand eczema"[Text Word] OR "hand dermatitis"[Text Word]	11,352
Drugs		
#2	dupilumab OR lebrikizumab OR nemolizumab or alitretinoin OR tacrolimus OR pimecrolimus OR ruxolitinib OR crisaborole OR desonide OR cyclosporine	23,858
Symptoms/Impact		
#3	symptom* OR pruritus OR itch* OR scratch* OR dry OR scaly OR crack* OR thick* OR leak* OR crust* OR swollen OR red OR broken OR bleed* OR HRQOL OR HRQL OR QOL OR "quality of life" OR sleep OR pain OR impact OR burden OR function OR functioning OR depress* OR anxiety OR anxious OR mood OR patient-reported or self-reported OR "patient report" OR "self report"	5,764,020
Study Type		
#4	"clinical trial" OR random* OR double-blind OR double-blinded OR control* OR placebo OR "Clinical Trial"[Publication Type]	2,563,879
Combination Search		
#5	#1 AND #2 AND #3 AND # 4	**260**

Search Number	Search Terms	Number of Results
Exclusionary Searches		
#6	“Animals”[Mesh] NOT “Humans”[Mesh]	1,143,792
#7	dog or dogs or canine or canines	69,265
#8	“Comment”[Publication Type] OR “Letter”[Publication Type] OR “Editorial”[Publication Type]	696,116
#9	**(#9) NOT (#10 OR #11)**	**213**

Note: Limitations: Publication date 2006 to 2017; English language. Search conducted March 7, 2017.

MeSH = Medical Subject Heading.

Online Resource 2. PROMs in Clinical Trials of AD Drugs

Study Reference	Study Design/ Patient Population	Intervention	PROMs Used	Findings Based on PRO Endpoints
Beck et al., 2014	Phase 2a RCT with 207 adults who had moderate-to-severe AD despite treatment with topical glucocorticoids and calcineurin inhibitors	Dupilumab v. placebo	Pruritus NRS (0-10, with higher scores indicating greater itch) 5-D pruritus scale (on which scores range from 5 to 25, with higher scores indicating greater itch)	Pruritus scores on the NRS decreased (indicating a reduction in itch) by 55.7% in the dupilumab group versus 15.1% in the placebo group ($P < 0.001$) at 12 weeks. 5-D pruritus results not reported.
Simpson et al., 2016b	Phase 2b RCT with 380 adults with moderate-to-severe AD	Dupilumab v. placebo	Pruritus NRS POEM HADS DLQI EQ-5D	Dupilumab reduced peak itch at 16 weeks relative to placebo by 1.1 to 3.2 points on Numeric Rating Scale ($P < 0.0001$ all doses, except 100 mg every 4 weeks $P < 0.05$); improved sleep and HRQOL on DLQI and EQ-5D ($P < .05$ all doses, except 100 mg every 4 weeks); and reduced anxiety and depression symptoms ($P < .05$ all doses).
Simpson et al., 2016a	Phase 3 Two 16-week studies Study 1 (SOLO 1) and Study 2 (SOLO 2): dupilumab (QW, Q2W) vs. placebo	Dupilumab v. placebo	Peak Pruritus NRS (key secondary endpoint) POEM HADS DLQI	<u>Combined results for SOLO 1 and SOLO 2</u> Peak Pruritus NRS: Scores improved at least 3 points or at least 4 points in significantly more patients receiving dupilumab v. placebo ($P < 0.001$ for all comparisons). By week 2, patient-reported scores with respect to itching were significantly better among patients receiving dupilumab than among those receiving placebo.

Study Reference	Study Design/ Patient Population	Intervention	PROMs Used	Findings Based on PRO Endpoints
	Study 1: 671 adults with moderate-to-severe AD Study 2: 708 adults with moderate-to-severe AD			POEM: Differences in patient-reported symptoms of AD and effect on sleep were significantly better for the dupilumab groups ($P < 0.001$). HADS: Among patients who had symptoms of anxiety or depression (HADS-A or HADS-D score ≥ 8) at baseline, significantly more dupilumab-treated patients than those receiving placebo had HADS-A and HADS-D scores of less than 8 at week 16 ($P < 0.001$). DLQI: Significantly more patients in the two dupilumab groups v. placebo groups had a reduction of at least 4 points (considered to be MCID) in total score ($P < .001$).
Ruzicka et al., 2017	Phase 2 12-week RCT with 264 adults with moderate-to-severe AD	Nemolizumab v. placebo	Pruritus VAS Pruritus VRS (0-4, none to very severe) Sleep disturbance VAS DLQI	At 12 weeks, there was a significant, dose-dependent reduction in pruritus VAS scores for the nemolizumab group versus placebo (P values ranged from 0.002 to < 0.001). Change in the score on the pruritus VRS were −36.8 ± 4.6% with 0.1 mg per kilogram, −50.9 ± 4.6% with 0.5 mg per kilogram, and −57.6 ± 4.6% with 2.0 mg per kilogram, as compared with −16.2 ± 5.0% with placebo (P values not reported). Changes in sleep disturbance VAS were −52.3 ± 5.8% with 0.1 mg of nemolizumab per kilogram, −59.1 ± 5.8% with 0.5 mg per kilogram, and −62.6 ± 5.9% with 2.0 mg

Study Reference	Study Design/ Patient Population	Intervention	PROMs Used	Findings Based on PRO Endpoints
				per kilogram, as compared with −31.9 ± 6.3% with placebo (*P* values not reported). The LS mean DLQI scores changed by –5 to –7 points for all nemolizumab Q4W groups (the criteria for MCID is > 4 units). At week 12, the LS mean change in DLQI score was –5.2, –6.1, and –7.0 in the nemolizumab 0.1, 0.5, and 2.0 mg/kg dose groups, respectively v. –4.3 for placebo.
Poole et al., 2009	Phase 3 Data from a clinical trial with 257 patients (16 and older) with mild, moderate, and severe AD	Tacrolimus v. placebo	SF-36	The active tacrolimus treatment group demonstrated improvement across all eight domains of the SF-36.
Reitamo and Allsopp, 2010	Phase 3 2 RCTs with 322 adults and children with mild, moderate, or severe AD	Tacrolimus v. placebo	Pruritus VAS DLQI/CDLQI	In adults, QOL improved from a score of 10.6 ± 6.7 to 5.5 ± 5.2 on the DLQI. Pruritus results not reported.
Doss et al., 2009	Phase 4 Patients 16 and older with moderate-to-severe AD of the face in whom conventional treatment	Tacrolimus v. fluticasone	Pruritus VAS	Baseline facial pruritus was severe in both groups (mean VAS scores 65 ± 25 mm in both groups). At day 7, median relative changes from baseline were –69.5% in the tacrolimus group and –71.8% in the fluticasone group, with further improvement at day 21 or EOT (–

Study Reference	Study Design/ Patient Population	Intervention	PROMs Used	Findings Based on PRO Endpoints
	was ineffectual or poorly tolerated			89.3% and –88.3%, respectively). Between-group differences were not statistically significant.
Kim and Kono, 2011	Pooled analysis of studies in eight Asian areas; total of 860 adults and children with AD	Tacrolimus	Pruritus VAS DLQI/CDLQI	The change in patient's assessment of pruritus measured by VAS score from baseline was -3 cm at the end of treatment. The improvements in total QOL score from baseline to the end of treatment were observed in adults and children.
Boguniewicz et al., 2007	Open-label, non-comparative, multicenter study Adult (n = 18) and pediatric (5-16 years, n = 22) patients with moderate-to-severe AD	Tacrolimus ointment (open-label, non-comparative)	SF-12 DLQI (CDQLI for pediatric patients) Bergner Physical Appearance Scale (measures patient perception of appearance) Productivity/ absenteeism	Adult patient DLQI scores showed downward trends, but they were not statistically significant (*P* = 1.38; *P* = .202). SF-12 scores showed mixed trends for physical and mental components; none were statistically significant. No statistically significant changes in patient or caregiver-reported physical appearance There was a significant decrease in work absenteeism from baseline (30.3% missing some work) to month 6 (0% missing some work).
Reitamo et al., 2008	Long-term follow-up study 4-year follow-up study with 782 AD patients aged 2 and older	Tacrolimus	Treatment satisfaction (excellent, very good, good, fair, and poor)	At the end of the study or at the time of withdrawal, 75.0% of the patients rated their satisfaction with treatment as excellent, very good or good.

Study Reference	Study Design/ Patient Population	Intervention	PROMs Used	Findings Based on PRO Endpoints
Leung et al., 2009	No phase specified RCT with 73 patients with AD, aged 2-49, with a clinical insensitivity to topical corticosteroids (determined by *Staphylococcus aureus* colonization and production of superantigens)	Pimecrolimus v. placebo	Pruritus (4-point scale, where 0=absent and 3=severe) Patient assessment of disease control (4-point scale, where 0=complete control and 3=uncontrolled disease)	An increase in *S. aureus* counts correlated with clinical worsening (week 6 vs. baseline) when assessed by IGA and patient assessments of pruritus severity and disease control. Patient assessments of disease control and pruritus were comparable between treatment groups at week 6.
Luger et al., 2007	Patient self-observation study in 3,502 patients with AD	Pimecrolimus	Pruritus (6-point scale) Erythema (6-point scale)	Patients reported marked improvement after only 2 weeks. Compared with the values recorded at treatment initiation from the previous year, the intensity of redness and pruritus decreased by nearly 2 points (from severe to mild).
Onumah and Kircik, 2013	Open-label patient preference study with 20 children (2 years and older) and adults with moderate AD	Pimecrolimus v. tacrolimus	Product preference questionnaire (included product rating/assessment, pruritus, burning, stinging, pain, and global assessment) DLQI/CDLQI	Mean change in DLQI was 4.7 for pimecrolimus compared with 3.6 for tacrolimus. Pimecrolimus scores were higher (i.e., better) in every category of the product preference questionnaire compared with tacrolimus.

Study Reference	Study Design/ Patient Population	Intervention	PROMs Used	Findings Based on PRO Endpoints
Kircik, 2014	Pilot study Open-label study in 20 AD patients aged 7 or older	Desonide hydrogel	Pruritus VAS	All subjects had 50% or greater improvement in pruritus at day 7.
Trookman et al., 2011	Patient preference study with 22 adults with mild-to-moderate AD	Desonide hydrogel	Patient preference	Desonide hydrogel rated highly in aesthetic attributes important to AD patients and was preferred by a majority of patients to other vehicles used in the past.
Trookman and Rizer, 2011	Patient preference study with 46 individuals 12 years of age and older with mild-to-moderate AD	Desonide hydrogel v. desonide ointment	Burning/stinging (0-3 scale) Dryness (0-3 scale) Pruritus (0-3 scale) Vehicle Preference Questionnaire	All symptom measures declined significantly from baseline at weeks 2 and 4 for patients receiving both desonide hydrogel and ointment ($P < 0.05$). Desonide hydrogel was rated significantly better than desonide ointment ($P < 0.05$) on absorption (week 4) and the (lack of) greasiness of the formulation (week 2).
Koppelhus et al., 2014	Randomized crossover study with 20 patients with severe AD	Cyclosporine A v. extracorporeal photopheresis	Pruritus VAS Patient global assessment	The average reduction in pruritus was a little higher for photopheresis treatment compared with cyclosporine, but the differences did not reach statistical significance. Photopheresis was rated “good” or “very good” by 74% of participants, while only 6% gave this rating to cyclosporine.
Paller et al., 2016	Phase 3 RCT with children (2 years and older) and	Crisaborole v. placebo	Pruritus severity (measure not specified)	Crisaborole-treated patients achieved improvement in pruritus earlier than vehicle-treated patients (pooled data, 1.37 vs 1.70 days, $P = 0.001$).

Study Reference	Study Design/ Patient Population	Intervention	PROMs Used	Findings Based on PRO Endpoints
	adults with mild-to-moderate AD			

AD = atopic dermatitis; CDLQI = Children's Dermatology Life Quality Index; DLQI = Dermatology Life Quality Index; EOT = end of treatment; EQ-5D = EuroQol 5-Dimensions; HADS = Hospital Anxiety and Depression Scale; HRQOL = health-related quality of life; IGA = Investigator Global Assessment; LS = least squares; MCID = minimal clinically important difference; NRS = numerical rating scale; POEM = Patient-Oriented Eczema Measure; PROM = patient-reported outcome measure; Q2W = every 2 weeks; Q4W = every 4 weeks; QW = every week; QOL = quality of life; RCT = randomized controlled trial; SF-36 = 36-Item Short Form Health Survey; VAS = visual analog scale; VRS = verbal rating scale.

Online Resource 3. PROMs in Clinical Trials of CHE Drugs

Study Reference	Study Design/ Patient Population	Intervention	PROMs Used	Findings Based on PRO Endpoints
Ruzicka et al., 2004	Phase 2 12-week RCT conducted in 43 outpatient clinics in Europe with 319 patients with moderate or severe refractory CHE	Oral alitretinoin 10 mg/d, 20 mg/d, and 40 mg/d, compared with placebo control	Patient global assessment of improvement DLQI	▪ Patient global assessment: Number of patients rating their response as "clear or almost clear" was significantly higher with all doses of alitretinoin than with placebo (P = .01 for 10 mg/d, P = .002 for 20 mg/d, and P < .001 for 40 mg/d) ▪ DLQI: Between-group differences were not statistically significant.
Fowler et al, 2014	Phase 3 RCT with 596 patients with severe CHE refractory to topical corticosteroids	Alitretinoin v. placebo	Patient global assessment Skindex-29	▪ A greater proportion of alitretinoin patients achieved patient global assessment of cleared or almost cleared at end of trial (OR = 4.05, P < 0.001). ▪ Skindex-29 total and subscale (Emotions, Symptoms, Functioning) scores were significantly improved at EOT for alitretinoin patients vs. placebo (P < 0.001 for all).
Ruzicka et al., 2008	24-week RCT conducted in 111 dermatology outpatient clinics in Europe and Canada with a total of 1,032 patients with severe refractory CHE	Oral alitretinoin 10 mg or 30 mg or placebo daily (1:2:2 ratio of placebo, 10 mg, 30 mg)	Patient global assessment of improvement (secondary endpoint)	Patient global assessment, n (%) clear or almost clear ▪ 30 mg 163 (40%) (P < 0.001 vs. placebo) ▪ 10 mg 101 (24%) (P < 0.02 vs. placebo) ▪ Placebo 31 (15%)

Study Reference	Study Design/ Patient Population	Intervention	PROMs Used	Findings Based on PRO Endpoints
Dirschka et al., 2011	Open-label study with 249 patients aged 18–75 years with severe CHE unresponsive to treatment with topical corticosteroids	Alitretinoin	Patient global assessment Pain VAS Pruritus VAS Categorical rating scale for pruritus PBI-HE	▪ Patient global assessment : 115 (46.2%) patients rated their disease as “clear or almost clear” at the end of treatment; results of the patient global assessment were “comparable with” PGA results (46.6% were responders, with PGA ratings of “clear” or “almost clear” hands, and 63.9% were classified as at least partial responders, with ratings of “clear,” “almost clear,” or “mild disease”). ▪ Pruritus categorical scale: At baseline, the intensity of pruritus was described as moderate (36.9% of patients) or severe (39.0% of patients), whereas only 4.0% described pruritus as absent. At end of treatment, pruritus was absent in 57.0% of the patients and 19.6% described their pruritus as severe (8.0%) or moderate (11.6%). ▪ VAS ratings of both pain and pruritus intensity showed corresponding decreases at the end of treatment (–33.2% and –49.6% mean change, and –88.9% and -89.4% median change from baseline, respectively) (*P* values not reported).
Hordinsky et al, 2010	RCT with 652 adults with mild-to-moderate CHE	Pimecrolimus v. placebo	Pruritus severity Burning severity Both measured on a 0-3 scale, where 0 = absent and 3 = severe	▪ The proportion of patients experiencing pruritus relief was significantly higher in the pimecrolimus group (83.7%) compared with the vehicle group (72.8%) at the end of week 6. ▪ No statistically significant difference was seen between the treatment groups with respect to burning.

CHE = chronic hand eczema; PBI-HE = Patient Benefit Index-Hand Eczema; RCT = randomized controlled trial; VAS = visual analog scale.