Evidence Review Group's Report

Title: Alitretinoin for the treatment of severe chronic hand eczema

Produced by	*Centre for Reviews and Dissemination (CRD), Centre for Health Economics (CHE)*
Authors	*Mike Paulden, Research Fellow, CHE* *Mark Rodgers, Research Fellow, CRD* *Susan Griffin, Research Fellow, CHE* *Russell Slack, Research Fellow, CRD* *Steven Duffy, Information Specialist, CRD* *John R Ingram, UK Dermatology Clinical Trials Network, University Hospital of Wales, Cardiff* *Nerys Woolacott, Senior Research Fellow, CRD* *Mark Sculpher, Professor of Health Economics, CHE*
Correspondence to	*Mike Paulden, Research Fellow, CHE*
Date completed	*2nd March 2009*

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 08/87/01.

Declared competing interests of the authors
None

Acknowledgements

Professor Hywel Williams, Professor of Dermato-Epidemiology, Queen's Medical Centre, Nottingham

Rider on responsibility for report
The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Table of contents

List of tables and figures

Glossary and list of abbreviations

CHE: Chronic Hand Eczema

DLQI: Dermatology Quality of Life Index

mTLSS: Modified Total Lesion Symptom Score

PaGA: Patient's Global Assessment

PGA: Physician's Global Assessment

PUVA: Psoralen and UVA treatment

Chronic hand eczema: A persistent, relapsing inflammatory condition of the skin, confined largely or wholly to the hands, characterised by thick, scaly skin that commonly gives rise to itchy blisters, redness, swelling and painful cracks in the skin.

Hyperkeratotic CHE: Hand eczema predominately characterized by thickening and hardening of the skin.

Pompholyx: A type of eczema characterised by itchy blisters, followed by inflamed and dry skin.

Note on use of page numbers

All page numbers given in parentheses in this ERG report refer to the manufacturer's original submission, unless otherwise stated. References to the ERG report are given in terms of section number (e.g. "see section 4.2.2 for details").

1 Summary

1.1 Scope of the submission

This report presents the ERG's assessment of the manufacturer's (Basilea Pharmaceuticals Limited) submission to NICE on the use of alitretinoin for the treatment of severe chronic hand eczema (CHE). The report includes an assessment of both the clinical and cost effectiveness evidence submitted by the company. The manufacturer's submission adhered to the scope for the appraisal issued by NICE in that it considered the use of alitretinoin (within the context of the licensed indication) in adults with severe chronic hand eczema refractory to topical steroid treatment and attempted to compare it with the stated comparators: PUVA, ciclosporin and azathioprine.

1.2 Summary of submitted clinical effectiveness evidence

The main clinical effectiveness data were derived from a single placebo controlled randomised controlled trial (RCT) of daily treatment with alitretinoin for 12 to 24 weeks, with follow-up for a further 24 weeks, in patients with severe CHE unresponsive to topical steroids. The results showed a statistically significantly greater proportion of patients achieved the primary endpoint of clear or almost clear hands (as assessed by the physician's global assessment (PGA)) by week 24 than did with placebo: 48% with alitretinoin 30 mg ($p<0.001$); 28% with alitretinoin 10 mg ($p<0.005$); 16.6% with placebo. The severity of disease was also reduced when assessed by patients using the patient global assessment (PaGA). The majority of patients who responded to alitretinoin treatment remained in remission during the 24 week follow up period (35% for 30mg, 28% for 10mg). A high PGA response rate to retreatment with alitretinoin was observed, though a similarly high response to placebo was observed among first-line 'placebo responders', and PGA results were not consistent with the PaGA evaluations. Dose-dependent headache was the most commonly reported adverse event in patients treated with alitretinoin, with rates of 20% in the alitretinoin 30 mg group and 11% of the alitretinoin 10mg group, respectively. Serious adverse events were rare, but alitretinoin was associated with increases in both total cholesterol and triglycerides, which has implications for risks of future cardiovascular events.

No direct comparisons of alitretinoin with any of the relevant treatment comparators (PUVA, ciclosporin or azathioprine) were available. Nor were any trial data on these comparators available to permit formal indirect comparisons of alitretinoin with its comparators.

1.3 Summary of submitted cost effectiveness evidence

The manufacturer's searches did not recover any existing economic evaluations of alitretinoin for the treatment of CHE, and so the manufacturer developed a *de novo* decision analytic model to estimate, over a time horizon of three years, the cost-effectiveness of alitretinoin versus the other relevant comparators identified by NICE. In response to the points of clarification put to them by the ERG regarding the initial submission, the manufacturer provided additional evidence and a revised decision analytic model.

In the manufacturer's original submission to NICE, the base case ICERs reported for alitretinoin were £8614 per QALY versus ciclosporin, -£469 per QALY versus PUVA (with alitretinoin dominant) and £10,612 per QALY versus azathioprine. These ICERs decreased as the time horizon was extended up to 20 years in sensitivity analyses. In patients with hyperkeratotic CHE and in women of child-bearing potential, the ICER increased slightly but remained below £20,000. When the utility values used in the model were replaced with those derived from an alternative study, these ICERs increased significantly (to £22,312 per QALY for alitretinoin versus azathioprine).

1.4 Commentary on the robustness of submitted evidence

1.4.1 Strengths

The manufacturer's submission incorporated a full systematic review of the literature of the effects of alitretinoin in severe CHE refractory to topical steroid treatment. The main findings are derived from a single generally well-conducted placebo controlled RCT and an associated follow-up trial of retreatment.

The submission also included a review of the literature of the cost-effectiveness of alitretinoin in severe CHE. As no existing economic evaluations were identified, the manufacturer undertook a *de novo* economic evaluation in order to compare alitretinoin with comparators identified by NICE, consisting of ciclosporin, PUVA and

azathioprine. The model estimated costs and quality-adjusted life years (QALYs) from the perspective of the NHS and PSS, which is consistent with NICE guidelines.

1.4.2 Weaknesses

At present, there is a relatively limited quantity of evidence available on the clinical effects of alitretinoin. Though the RCTs presented were adequately designed and conducted, the ERG noted high numbers of withdrawals from the main efficacy trial, a lack of clear evidence for the reported subgroup effects and unexplained inconsistencies between PGA and PaGA scores in the retreatment trial.

Limitations in the submitted evidence primarily impacted on the generalisability of the manufacturer's conclusions to clinical practice. The main observed effects of alitretinoin were relative to placebo with additional emollients where required. Therefore it remains unknown to what extent alitretinoin is effective relative to emollients and topical corticosteroids combined (the current first-line treatment choice).

For inclusion in the main RCT (BAP00089), diagnosis as "severe" on the Physician's Global Assessment (PGA) outcome measure was a prerequisite. In clinical practice, it seems likely that a proportion of patients considered for treatment with alitretinoin would fall into the less severe PGA "moderate" state. There is some evidence from the phase II trial BAP00003 that a 'PGA moderate' CHE population will respond to alitretinoin treatment but there is no evidence for the effects of the licensed 30mg dose in this population.

The cost-effectiveness section of the submission had major shortcomings. The efficacy estimates for treatments other than alitretinoin were based on expert clinical opinion only. While the use of expert opinion may be justified where trial data do not exist to inform the relevant parameters, it should be elicited in a methodologically rigorous manner. The ERG remains unconvinced that this elicitation process generated reliable estimates of the efficacy of each of the comparator treatments. The estimates of HRQL were derived in a two-stage prediction model that incorporated an algorithm developed in patients with psoriasis.

Serious issues remain around the implementation of the model in Excel. Inspection of the VBA code indicated that a number of the assumptions given in the written submission were not implemented correctly. In particular, the first four weeks of

every subsequent treatment cycle were omitted. The definition of relapse used in the model did not correspond to that used in the relevant clinical trials. As a consequence the estimated costs and health outcomes presented by the manufacturer may be regarded as unreliable. The ERG attempted to amend the model to provide more appropriate estimates of the ICERs but in some cases this was not feasible.

Furthermore, the model originally submitted to NICE did not include a "supportive care" (or "placebo") arm and the treatment effects for alitretinoin were not placebo adjusted; as such, the model did not address whether alitretinoin was a cost-effective alternative to supportive care. Consequently, the ERG does not regard the ICERs generated by the manufacturer's original model as providing a reliable indication of the cost-effectiveness of alitretinoin compared to each of the comparators considered.

1.4.3 Areas of uncertainty

Crucially, there is no evidence on the efficacy and safety of alitretinoin beyond around 48 weeks. Given the chronic recurring nature of CHE, longer term follow-up is required to detect potentially rare adverse events and possibly to characterise the cardiovascular risks posed by the observed increase in cholesterol levels associated with alitretinoin treatment.

There was also no direct or indirect evidence presented for the clinical effects of alitretinoin relative to the comparators specified in the scope for the treatment of CHE (PUVA, ciclosporin and azathioprine). No additional evidence was identified by the ERG.

A change in threshold for the definition of 'relapse' from 75% to 50% of baseline mTLSS substantially reduced the time to relapse observed in the 30 mg alitretinoin group. If clinicians were to consider retreatment for less severe 'relapses', this would have clinical and cost implications in terms of the reduced the between treatment periods.

As the relief of symptoms and consequent improvement in health related quality of life is the aim of treatment for chronic hand eczema, the ERG believes that the economic evaluation of alitretinoin should be based on good evidence of the improvement in health related quality of life offered by alitretinoin. However, the

estimates used in the submission are subject to a great deal of uncertainty due to the two-stage prediction employed and the paucity of direct observations in the population of interest.

The ERG modified the manufacturer's model to examine the impact of altering some of the key assumptions. However, as the manufacturer did not undertake a probabilistic sensitivity analysis, the combined impact of uncertainty in the inputs to the economic model on the overall decision uncertainty could not be evaluated.

1.5 Key issues

Longer-term follow-up of trials or the implementation of registries are required to better establish the longer term efficacy or safety of alitretinoin. Future studies should include a relevant HRQL measure (such as the DLQI and EQ5D) alongside measures of therapeutic response and may want to establish the efficacy of alitretinoin relative to current first-line treatment (emollients plus topical steroids) and other treatments which are used in this indication (PUVA, azathioprine, ciclosporin).

In response to a request from the ERG, the manufacturer provided a revised model with a "placebo" arm, and the comparison of alitretinoin with placebo made in this revised model is of greater merit given the more reliable efficacy data in the comparator arm. In this analysis, alitretinoin was reported to have an ICER of £12,931 per QALY gained versus supportive care (placebo). However, the omission of adverse events entirely from this revised model, in combination with a number of other factors, means that the model underestimates the costs of treatment associated with alitretinoin and so the true ICER may be higher.

The manufacturer assumed that patients receiving alitretinoin visited the dermatologist every four weeks and ceased treatment as soon as they responded, even if this was after only four or eight weeks of treatment. If in practice patients would receive treatment for longer than this then the manufacturer's model will have significantly underestimated the costs to the NHS.

Additional analyses undertaken by the ERG produced ICERs close to £30,000 per QALY gained for alitretinoin versus supportive care. However, there remains considerable uncertainty as to the true ICER of alitretinoin versus the relevant treatment comparators.

2 Background

2.1 Critique of manufacturer's description of underlying health problem

The manufacturer provides an accurate summary of severe chronic hand eczema (CHE), largely based on Diepgen et al[1] and Agner et al.[2]

2.2 Critique of manufacturer's overview of current service provision

In general, the manufacturer provides a reasonable overview, though the emphasis of a certain specific points could be questioned:

- Page 13, paragraph 6, the authors state that the efficacy of topical immunomodulators is "very low" in severe CHE. The clinical advisor consulted by the ERG suggested that the efficacy of 0.1% tacrolimus ointment is probably equivalent to that of a potent topical steroid. In the same paragraph, the manufacturer states that "topical immunomodulators are associated with a risk of cancer" – this remains controversial and most of the evidence suggests that there is not a significantly increased risk.

- Page 13, paragraph 8. The ERG's clinical advisor considers "pronounced toxicity" of ciclosporin and azathioprine to be an overstatement. When used appropriately, it would be more correct to state they have "potential toxicity". The manufacturer goes on to state that potential toxicity "limits (the use of ciclosporin) on any long-term basis" – current British Association of Dermatology guidelines for ciclosporin in psoriasis permit long-term treatment of a sub-group of severely affected patients. However, the guidelines recommend that intermittent courses are preferable to continuous treatment, especially for treatment duration greater than two years. There is no specific guidance for CHE.

3 Critique of manufacturer's definition of decision problem

3.1 Population

The manufacturer states (p.4) that the population of interest is "adults with severe chronic eczema of the hand that is unresponsive to topical corticosteroids". This reflects the population specified in the final scope issued by NICE. Of the included studies, one phase III trial included only patients classified as 'severe' according to the Physician's Global Assessment scale (PGA; see section 4.1.6), and one phase II trial included patients classified as either 'moderate' or 'severe'. In practice, the most important factor would be the impact of the condition on the patient's activities, usually quantified by a Patient Global Assessment (PGA) or a quality of life score such as the Dermatology Life Quality Index (DLQI). Hence in clinical practice patients qualifying for treatment with alitretinoin may equate to either 'severe' or 'moderate' states as defined by the PGA alone.

3.2 Intervention

The NICE final scope indicates this to be alitretinoin.

The manufacturer specifies that use of alitretinoin in its licensed indication, which is 10 mg - 30 mg once daily for 12 to 24 weeks depending on response, with discontinuation of therapy to be considered for patients who still have severe disease after the initial 12 weeks of treatment (p3-4). The Summary of Product Characteristics (SPC) states that, in the event of relapse, patients may benefit from further treatment courses of alitretinoin.

3.3 Comparators

The decision problem addressed in the manufacturer's submission specified relevant comparators to be ciclosporin, azathioprine, and oral and topical PUVA. This reflects the final scope issued by NICE.

However, after discussion with the ERG's clinical advisor, and given that alitretinoin is the only licensed treatment for severe CHE, the ERG requested that the manufacturer also include in their model an appropriate 'supportive care' arm, which could include ongoing treatment with topical steroids. In the manufacturer's revised model, alitretinoin was compared to placebo, with both groups receiving supportive treatments in the form of emollients and dermatologist visits, but not steroids. The

manufacturer's response is discussed in Section 4.2.1.1 and full details of the revised modelling are given in Chapter 5.

3.4 Outcomes

The outcomes considered in the manufacturer's submission reflected those specified in the scope: disease severity (Physicians Global Assessment (PGA)) and Patient's Global Assessment (PaGA)), symptom control (modified total lesion symptom score (mTLSS)), disease free period/maintenance of remission, time to relapse, adverse effects and health related quality of life (Dermatology Quality of Life Index (DLQI)) (p35-37).

Relapse was defined as an mTLSS score of 75% of the baseline value (p36). Since patients with mTLSS scores just less than 75% are still likely to have severe CHE, the ERG asked the manufacturer to consider the impact of defining relapse with different mTLSS thresholds.

The direct effect of treatment on DLQI was not reported. Instead the manufacturer explored the relationship between DLQI and change in PGA state, independent of treatment effect, based on the findings of a phase II trial (BAP00003). The manufacturer states that this phase II study was underpowered to detect a treatment effect on DLQI. The ERG requested that the manufacturer provide full details of the DLQI results from this study.

3.5 Time frame

In the included studies, patients received alitretinoin for 12 or 24 weeks depending upon response (the manufacturer clarified that just 3 patients were allowed to discontinue for "early improvement" at weeks 16 or 20). Responding patients were monitored for relapse for up to 24 treatment-free weeks. There does not appear to be direct clinical evidence on the effectiveness and safety of alitretinoin beyond 48 weeks. In fact, given the very high rate of withdrawals (25.5% in BAP00089) and the fact that patients responding at 12 weeks were observed for a maximum of 36 weeks, the average time over which patients were actually observed is likely to be considerably shorter than 48 weeks.

3.6 Other relevant factors

N/A

4 Clinical Effectiveness

4.1 Critique of manufacturer's approach

4.1.1 Description of manufacturer's search strategy and comment on whether the search strategy was appropriate.

The manufacturer's submission described the search strategy used to identify published trials of alitretinoin for chronic hand eczema and stated clearly that full details of the search strategies used were reported in the appendices (6.1. Identification of studies). It explained that the strategies were adapted from those used in a Cochrane review protocol.[3] A simple flowchart giving an overview of the literature search was presented. The submission also made clear that this strategy was used to identify comparator studies (RCTs of PUVA, ciclosporin and azathioprine for CHE).

The brief description of the search strategy used to identify the economic literature in the cost effectiveness section of the report was also sufficient, and once again clearly pointed to the appendices for full details of the search methodology.

A detailed review of the search strategies employed identified a number of flaws/potential limitations (Appendix 1). However, despite the issues raised after examination of the strategies, it is unlikely that the company failed to identify the existing trials of alitretinoin. The only trials were those of the manufacturer and they had all relevant data. However, it is possible that potentially useful comparator trials were missed. Three of the included PUVA studies would not have been identified by the searches used; others may have been missed.

Similarly, the searches for economic evaluations were adequate if over sensitive (Appendix 1). However, It is not known how studies with QoL data, utilities, safety data, cost data, etc., were identified when no additional searches were conducted.

The ERG completed searches of its own to take into account some of the issues raised in its review of the manufacturer's search strategies. The ERG search strategies were based on those used in the submission, but with some amendments: MEDLINE, MEDLINE In-Process and EMBASE were searched separately so that both MeSH and Emtree terms were recognised (and to account for the lack of indexing in MEDLINE In-Process); PUVA free text terms were added; the 'treat$ or

therap$' search line was discounted; CENTRAL was searched; the term 'effective$' was removed from the economic strategy.

The results were then deduplicated against the submission strategy results. An additional 524 records were retrieved for the clinical evidence search. Only 21 additional records were retrieved from the economic searches. No clinical studies beyond those included in the manufacturer's submission were identified as being potentially worth inclusion.

See Appendix 1 for the strategies used.

4.1.2 Statement of the inclusion/exclusion criteria used in the study selection and comment on whether they were appropriate.

The systematic review of clinical effectiveness initially included RCTs that compared alitretinoin with a comparator (including placebo) for the treatment of CHE. Though not stated as an inclusion criterion, all included RCTs evaluating alitretinoin were conducted in patients with CHE unresponsive to topical steroids. The broad intervention and participant criteria were appropriate and - though it is not stated whether search results were double-screened to prevent errors and selection bias – all the relevant published data on clinical effectiveness appear to have been captured.

The manufacturer also conducted an additional search to identify RCTs of PUVA, ciclosporin, or azathioprine for the treatment of CHE on the basis that none of the identified RCTs of alitretinoin used a comparator other than placebo. This appeared appropriate. Again, no attempts appear to have been made to minimise errors or bias in the selection process, but independent searching and screening (using the same criteria) by the ERG did not identify any extra relevant studies that were not accounted for in the submission.

Several uncontrolled studies of comparator interventions (particularly for PUVA) were excluded from the submission which was appropriate because uncontrolled studies such as this tend to provide exaggerated estimates of efficacy. It should be noted however, that additional efficacy and safety data from an uncontrolled open-label study of 30 mg alitretinoin (BAP00626) were also presented in the submission. Though the efficacy data from this study were appropriately presented separately

from the RCT data (and not used in the cost-effectiveness model), the likely unreliability of these data must be borne in mind.

4.1.3 Table of identified studies. What studies were included in the submission and what were excluded.

The review of clinical effectiveness and safety of alitretinoin in CHE comprised only Basilea-funded studies; primarily three: a phase II randomised trial comparing three doses of alitretinoin (10 mg, 20 mg, 40 mg) against placebo (BAP00003); a phase III RCT evaluating 10 mg and 30 mg daily doses of alitretinoin versus placebo (BAP00089); and an extension of trial BAP00089 in which non-responding and responding-relapsing patients were followed-up (BAP00091). In BAP00091, patients who responded to treatment during BAP00089 but relapsed during follow-up were randomised to receive placebo or the same dose of alitretinoin as previously. All patients who did not respond during BAP00089 were allocated to receive 30 mg alitretinoin daily.

Additional data from two other studies are briefly presented in the manufacturer's submission: a small RCT comparing 10 mg vs. 30 mg daily doses of alitretinoin (BAP00200) and a larger uncontrolled open-label study of 30 mg alitretinoin (BAP00626).

Of 13 identified studies evaluating the efficacy of PUVA in hand eczema, eight were considered for inclusion in the review of comparator studies. These largely compared topical or oral PUVA with UVA-1, UVB, radiotherapy or no treatment. One study compared oral PUVA at home against hospital bath PUVA. Reported treatment durations were between 3 weeks to 3 months. At the request of the ERG, the manufacturer provided the reasons for excluding five of the identified studies. The reasons provided were that three did not provide separate results for hand and foot, one was considered inadequately controlled, and the results of another were not adequately described in English (Appendix 2).

Ultimately, only four studies were included in the manufacturer's synthesis of PUVA studies; a further four were excluded because they reported only mean reduction in severity as an outcome. The ERG requested further details on all eight included PUVA studies, and these are presented in Section 4.2.2.

A single study evaluating ciclosporin compared 3 mg/kg/day ciclosporin against topical corticosteroid (betamethasone-17, 21-dipropionate) treatment for 6 weeks (see section 4.2.2 for details). No studies evaluating azathioprine in CHE were reported.

4.1.4 Details of any relevant studies that were not included in the submission?

The ERG's independent search of the literature did not retrieve any additional studies meeting the review inclusion criteria.

The manufacturer's submission mentions an ongoing open-label multi-centre study investigating the safety and efficacy of alitretinoin in relapsed CHE (BAP00731). It is unclear whether any interim data are available from this trial.

4.1.5 Description and critique of manufacturer's approach to validity assessment

The included placebo-controlled alitretinoin RCTs were described in terms of allocation concealment, randomisation technique, length of follow-up, blinding of outcome assessors, justification of sample size, parallel/crossover design, whether conducted in the UK, consistency of dosing regimes with the SPC, comparability of study groups, and appropriateness of statistical analysis. This approach was broadly adequate, though some minor details were inconsistently reported (e.g. the manufacturer's critical appraisal states that concealment was not broken at all during BAP00003, but does not state whether or not this was the case for BAP00089, and the validity of the RCT comparing different doses of alitretinoin without a placebo group (BAP00200) was not assessed).

Validity assessment of the single uncontrolled alitretinoin study was inadequate, being simply limited to a statement that "assessment of response may be influenced by the lack of blinding" and "the lack of control group permits limited objective assessment of treatment effects". For example, there was no discussion of participant characteristics, follow-up or measurement of potential confounders.

The validity of studies evaluating comparator interventions was described more briefly than for the alitretinoin studies, and was combined with brief data extraction in

table 6.6.2 of the manufacturer's submission. These tables gave a reasonable overview of the eight comparator studies considered eligible for the indirect treatment comparison. At the request of the ERG, the manufacturer provided details of the five additional PUVA studies that were included in the review but not summarised in the original submission. These were not validity assessed, presumably on the basis that that they were not used to inform the efficacy comparison.

4.1.6 Description and critique of manufacturers outcome selection

To a greater or lesser extent, the manufacturer's submission addresses each of the outcomes specified in the final scope issued by NICE (see description of each on p35-37).

The primary outcome measure used in the trials was severity of CHE according to the Physician's Global Assessment (PGA), which combined the grading of disease severity against a photographic guide, with an indication of symptoms (pruritis/pain) and degree of functional impairment. The PGA describes five states of CHE severity (clear, almost clear, mild, moderate and severe), of which the combined "clear/almost clear" category was used to define response to treatment in the included trials.

In addition to physician assessment of disease by PGA, the six-state patient global assessment (PaGA) was measured as a secondary endpoint (again, with the "clear/almost clear" states defined as remission). In clinical practice, it is important to verify that patients perceive similar benefit from treatment to their physicians, in order to ensure concordance with therapy that is invariably less closely supervised than in clinical trials. Correlation of PaGA with PGA in the phase III trial provides internal validation that the PGA measures outcomes of treatment that are meaningful for physicians and patients alike. The PGA includes one symptom (pruritus/pain) that cannot be linked to the photographic guide but is clearly relevant to disease severity and is a major driver for patients to seek medical help.

Symptom control was also measured using the continuous Modified Total Lesion Symptom Score (mTLSS), in which a four-point scale was used to grade seven different signs and symptoms of CHE. Though response/remission was primarily defined in terms of PGA state, relapse was not. Instead, the manufacturer's submission defines relapse as an mTLSS score 75% that of the baseline value. The manufacturers state that this figure was considered by dermatologists to reflect the

usual working definition of relapse “sufficient to require re-treatment with systemic agents or phototherapy”. However, clinical advice to the ERG indicated that 75% of baseline mTLSS score might be a rather high threshold for retreatment. At the request of the ERG, the manufacturer provided additional data on the influence of applying a less stringent definition (50% of original mTLSS) on time to relapse from trial BAP00089. This is presented and discussed further in section 4.2.2

Despite a clear focus of CHE treatment being to improve health-related quality of life, this outcome was only directly measured in the phase II trial BAP00003, using the Dermatology Life Quality Index (DLQI). The manufacturer clarified that health related quality of life was not measured in either the BAP00089 or BAP00091 studies. The main quality of life analysis presented in the manufacturer’s submission of clinical efficacy examined the relationship between PGA state and DLQI, independent of treatment effect, using data from BAP00003 and an unpublished study of DLQI in adults with CHE[4](see section 4.2.2).

In general, relevant safety outcomes appear to have been measured (p67-71). The manufacturer determined the safety of alitretinoin through an evaluation of adverse events, serious adverse events and discontinuation due to adverse events reported in the identified clinical trials (BAP00089, BAP00091, BAP00003 and BAP00200), with an emphasis on the phase III clinical trial (BAP00089). Frequent adverse events observed in an uncontrolled open label study (BAP00626) and laboratory changes (thyroid stimulating hormone, cholesterol, triglycerides etc) observed in BAP00089 were also reported. At the request of the ERG, the manufacturer provided full and complete details of adverse events leading to treatment discontinuation in both BAP00089 and BAP00091.

In addition, the manufacturer briefly discussed a number of ‘special safety assessments’ investigating the effects of alitretinoin on psychiatric status, ophthalmological health, skeletal abnormalities and bone mineral density. The ERG asked for full tabulated data for these analyses, which the manufacturer duly provided.

4.1.7 Describe and critique the statistical approach used

The main results for the 30 mg dose of alitretinoin were derived from a single RCT (BAP00089). The submission includes a meta-analysis of results (response rates)

from patients receiving 10 mg alitretinoin in trials BAP00003 and BAP00089. This comparison appropriately excluded data from trials BAP00091 (overlapping participants with BAP00089) and BAP00200 (no placebo comparison). The studies were pooled appropriately using a fixed-effects model, with an assessment of statistical heterogeneity.

Data were presented for response by subtype of CHE: predominantly hyperkeratotic, hyperkeratotic and pompholyx, and pompholyx alone. These data were derived from a single trial (BAP00089) and the distinction between the subtypes is unclear, with overlap between them. The evidence for there being any subgroup in which the efficacy of alitretinoin is greater than in the overall CHE population is weak.

The review of comparator studies (labelled 'indirect treatment comparison' in the submission, though it does not involve any formal statistical comparison or any comparison with alitretinoin) involved pooling response rates from four studies comparing PUVA against UVA/UVB/no treatment using a random effects model. However, given the substantial and clearly apparent clinical heterogeneity and highly significant statistical heterogeneity among these four studies, this pooling is unlikely to have been appropriate.

In fact, data from the identified comparator studies were not used to inform any subsequent analyses – the submission simply states that "the efficacy of comparators was informed by clinical opinion" for the economic evaluation. The ERG asked for further details of how estimates of parameters were elicited from clinical experts. The manufacturer indicated that a panel of six experts provided consensus estimates of the distribution of patients between the severe, moderate, mild, and clear/almost clear PGA states at 4 weekly intervals. Though the expert panel provided point estimates of comparator efficacy, they were not asked to provide estimates of the associated uncertainty (see section 5.2.1 for further critique of these methods).

In order to assess, to some degree, the reliability of the elicited responses rates, the ERG present the results of all 8 included PUVA studies, without attempting quantitative pooling (see section 4.2.2).

4.1.8 Summary statement

The submission clearly reflects the decision problem in terms of participants, and interventions. However, the clinical efficacy section does not provide any indication of the effects of alitretinoin relative to those of the comparators identified in the final scope based on empirical studies. The outcomes selected for the assessment of the efficacy of alitretinoin are relevant, though there may be some issues around the definition of relapse, and the health-related quality of life data presented in the submission is limited and indirect.

The set of included studies appears to be both relevant and complete, though more might have been made of data from the identified comparator studies.

4.2 Summary of submitted evidence

4.2.1 Summary of results

The manufacturer concludes that alitretinoin is an effective, convenient, once-daily oral therapy for the treatment of severe CHE unresponsive to topical corticosteroids. They state that most frequent adverse event was headache which was considered to be dose dependant. For laboratory parameters (p.69) there were increases in total cholesterol and triglycerides, with greater frequency in the 30 mg dose as opposed to the 10 mg group.

The manufacturer identified only very limited data regarding the efficacy of comparator treatments in CHE. A small number of PUVA trials were considered, all of which were small trials with varying populations and outcomes. There was a single small ciclosporin trial that also met the inclusion criteria for a comparator, though no controlled clinical data for azathioprine were identified. The comparator treatments had significant underlying heterogeneity (p.66).

4.2.2 Critique of submitted evidence syntheses

4.2.2.1 Efficacy of alitretinoin

Table 1 provides a summary of the findings from all included controlled trials of alitretinoin.

Table 1 Primary and secondary study endpoints from controlled trials included in manufacturer's submission

Trial	Treatment	Response: PGA[1] (95% CI)	Response: PaGA[1]	Symptom change: mTLSS[2] (95% CI)	Health related quality of life: DLQI[3]	Relapse rate[4]
BAP00089	Placebo	16.6% (11.8, 22.4)	15%	-39% (-47, -27)	-	-
	10 mg	27.5% (23.3, 32.1) P<0.005*	24% (p<0.02)*	-56% (-63, -50) (p<0.001)*	-	29.6% (at 6 months)
	30 mg	47.7% (42.7, 52.6) P<0.001*	40% (p<0.001)*	-75% (-79, -69) (p<0.001)*	-	37.4% (at 6 months)
BAP00091 (Cohort A)	Placebo (previously placebo)	69.2%	23.1%	-40.3%	-	-
	Placebo (previously 10mg)	10%	-	-	-	-
	Placebo (previously 30mg)	8.3%	-	-	-	-
	10 mg	47.6%	75.5%	-78.8% (p=0.02)**	-	-
	30 mg	79.6%	38.1%	-67.4% (p<0.001)**	-	-
BAP00091 (Cohort B)	30 mg	46.2%	42.4%	-49.7%	-	-
BAP00003	Placebo	27%	12%	-25% (-42, -14)	-2	26%
	10 mg	39% (p=ns)*	29% (p=0.014)*	-59% (-73, -42) (p=0.03)*	-2	25%
	20 mg	41% (p=ns)*	34% (p=0.002)*	-52% (-73, -42) (p=0.002)*	-3	26%
	40 mg	53% (p<0.001)*	43% (p<0.001)*	-59% (-80, -44) (p<0.001)*	-3	32.5%
BAP00200	10mg	12.5% (1.6, 38.3)	-	-	-	-
	30mg	62.5% (35.4, 84.8)	-	-	-	-

Note: [1] % with clear/almost clear hands; [2] Median change in mTLSS score from baseline; [3] Median within-patient change from baseline to week 12; [4] % with mTLSS score 75% of baseline value. * compared with placebo ** compared with placebo (previously 30 mg)

Validity of submission's primary source of evidence

The submission's findings draw heavily upon the BAP00089 trial, a phase III multicentre, well conducted placebo-controlled RCT that clearly reports randomisation generation, concealment of allocation and intention-to-treat analysis. The primary endpoint is unambiguous and the duration of the study (48 weeks) reflects the duration and treatment of a single episode of severe CHE, but is rather short for the assessment of a potentially life long intermittent therapy. The placebo arm of the trial, which required participants with severe CHE to receive only emollients, does not reflect frequent clinical practice where patients may continue to use topical steroids despite their limited benefits. However, it does reflect the population defined in the NICE final scope for the appraisal of alitretinoin (adult patients with severe CHE refractory to potent topical steroids), and eligibility for treatment in the BAP00089 trial depended on the documentation of no benefit or inadequate benefit from previous topical steroids or intolerance for these agents (either no response or only a transient response to at least 8 weeks of topical corticosteroid therapy including 4 weeks of treatment with super-potent agents such as clobetasol propionate). Furthermore, the manufacturer has stated that, as there are potential adverse effects associated with the long-term use of topical steroids, there is no clear justification for their further use in patients unlikely to derive any obvious benefit. The manufacturer also refers to reasonable rates of PGA-defined response in the placebo arm of trial BAP00089. They suggest this refutes the need for topical steroids in the "supportive care" of patients with severe CHE, provided the standard supportive care of CHE can be optimised. However, clinical advice to the ERG indicated that the addition of topical corticosteroids to supportive care would have almost certainly further improved the response rate in this arm.

Response rates to alitretinoin

This trial found that clear or almost clear skin was reported for 47.7% of patients within 12-24 weeks of treatment with 30 mg alitretinoin, compared with 27.5% for 30 mg alitretinoin and 16.6% for placebo ($p<0.001$ and $p<0.005$). A 75% median reduction in signs and symptoms of CHE, measured by mTLSS, was observed after 24 weeks in the 30 mg alitretinoin treatment group and a 56% improvement in the 10 mg alitretinoin treatment group (both $p<0.001$ compared to placebo) (see table 6.4.3 on p44 and table 6.4.4 on p46) .

A meta-analysis of response rates (clear/almost clear) with 10 mg alitretinoin from the two RCTs that reported this gave a pooled odds ratio of 1.89 (95% CI: 1.32, 2.72; p=0.0004), indicating a significantly better response rate with low dose alitretinoin than with placebo (p53). The claim of the investigators that alitretinoin induced remission in a high proportion of participants may be subjective, given that just under half responded to the 30 mg dose and 28% to the 10 mg dose. However, such rates may be considered 'high' in patients who have not responded to other available treatments. The observed response for 30mg alitretinoin appears to represent a significant advance given the limited therapy previously available, but again this is in comparison to emollients only.

Table 2 Summary of Withdrawals from BAP00089

	Placebo	10 mg	30 mg
Percentage of withdrawals	33% (n=68)	24% (n=99)	26% (n=106)
Main Reasons			
Insufficient response	20.5% (n=42)	8.4% (n=35)	7.8% (n=32)
Refused treatment/lack of co-operation	5.9% (n=12)	5.7% (n=24)	3.9% (n=16)
Adverse Events	5.4% (n=11)	5.7% (n=24)	9.5% (n=39)

One concern with BAP00089 is that there were substantial numbers of withdrawals from each arm of the study (see Table 2 above). Withdrawal due to adverse events (mostly headaches) appears to be most prevalent in the 30 mg alitretinoin arm whereas for placebo and 10 mg alitretinoin insufficient response appears to be the main factor. Athough the manufacturers correctly analysed these results on an intention-to-treat basis, the large numbers of withdrawals is noteworthy.

In the open-label multi-centre study of 249 patients with severe CHE refractory to topical steroids (BAP00626), response rates to 30mg alitretinoin were also consistent with those observed in the RCTs, with 46.6% of patients achieving clear or almost clear hands by PGA assessment.

Remission and relapse

After 24-weeks follow-up of responders in the BAP00089 study, during which no other active medication was permitted, 65% and 72% of patients who had received 30 mg and 10 mg alitretinoin respectively remained in remission, with a what appeared to be a median time to relapse of 168 days for 30 mg alitretinoin, 190 days

for 10 mg alitretinoin and 168 days for placebo (table 6.4.4 on p.46; N.B. the range values reported for the two alitretinioin groups appear to have been unintentionally switched and labelled as confidence intervals). The proportion of placebo patients in remission does not appear to have been reported. The claim that alitretinoin induced durable remission is not evident in the data which reports the same median time to relapse in both the alitretinoin 30 mg and placebo groups (p.46).

Again, patient withdrawal rates were notable. Of patients entered into BAP00089, a further 20.6% (seven patients) were withdrawn without 24-week post-treatment follow-up in the placebo arm, 23.5% (27 patients) in the 10 mg alitretinoin arm and 27.7% (58 patients) in the 30 mg alitretinoin arm (mainly due to unspecified "administrative reasons", see figure 6.3.2 of manufacturer's submission).

Impact of changing definition of relapse

After taking clinical advice, the ERG requested that the manufacturer provide additional data on the influence of applying a less stringent definition (50% of original mTLSS) of relapse from trial BAP00089. It can be seen from table 3 below that the change in threshold for relapse from 75% to 50% of baseline mTLSS appears to have little influence on median time to relapse in the placebo or 10 mg alitretinoin treatment groups, but does substantially reduce the time to relapse in the 30 mg alitretinoin group. The first quartile relapse rates are reduced for all three groups.

Table 3: Time to relapse (days) with different criteria for relapse (BAP00089)

		Placebo	10 mg	30 mg
Median	PGA Mild	86	63	56
	PGA Moderate	162	162	107
	PGA Severe	NA	NA	NA
	mTLSS 50%	165	190	99
	mTLSS 75%	168	190	168
1st Quartile	PGA Mild	29	30	29
	PGA Moderate	60	63	56
	PGA Severe	112	205	99
	mTLSS 50%	64	63	53
	mTLSS 75%	86	147	84

The manufacturer points out that the 75% mTLSS values fall between the moderate and severe PGA relapse rates, whereas the 50% mTLSS values are very close to PGA moderate. On this basis, they conclude that a less stringent relapse definition of 50% mTLSS would yield a population with moderately severe CHE who would unlikely to be immediately retreated with systemic therapy. However, it should be noted that this appears to contradict the manufacturer's earlier statement on p.26 of

their response to clarifications that “Basilea does not assume that diagnosis of severe chronic hand eczema requires the patient to be identified as in PGA category ‘severe’”. In fact, as indicated in Table4, 35.9% of all relapsing patients recruited for retreatment in trial BAP00091 were in the moderate PGA state. Clinical advice to the ERG suggests that if the patient has had benefit and not experienced adverse effects, then both they and their dermatologist are likely to consider re-treatment at lower levels of disease severity.

Table 4: PGA severity and mTLSS for BAP00091 study patients at baseline

	Cohort A Relapse in BAP00089			
		10 mg	**30 mg**	Placebo
Number of Patients (ITT)	21	49	47	
Physician's Global Assessment at Baseline				
Clear	0	0	0	
Almost Clear		1 (4.8%)	0	0
Mild Disease		1 (4.8%)	0	0
Moderate Disease		9 (42.9%)	**15 (30.6%)**	18 (38.3%)
Severe Disease		10 (47.6%)	**34 (69.4%)**	29 (61.7%)
• mTLSS at Baseline				
n	21	49	47	
Mean	12.6	13.3	13.4	
SD	3.19	2.36	2.35	
Median		12.0	13.0	14.0

Source: B91T09.sas 10may07

Analysis of 24-week response rates by subtypes of CHE in BAP00089 (see Table 5) suggested that patients with hyperkeratotic symptoms might respond better to alitretinoin than do other subtypes, though the was not powered to consider sub-groups. The ‘pompholyx only’ group in particular is very small. However, the potential trend observed fits with the effects of retinoids in other skin diseases.

Table 5: Subgroup response rates from BAP00089

CHE subtype (% of ITT population)	Hyperkeratotic (64%)	Hyperkeratotic/Pompholyx (22%)	Pompholyx (5%)
Clear/almost clear (PGA)	30mg: 54% 10 mg: 30% Placebo: 12%	30mg: 33% 10 mg: 23% Placebo: 12%	30mg: 33% 10 mg: 22% Placebo: 30%

Response to retreatment

Trial BAP00091 comprised two patient cohorts. Cohort A involved patients who had responded to alitretinoin with clear/almost clear hands and who subsequently relapsed within the 24-week follow-up period, whereas cohort B comprised patients who were not classed as responders in the BAP00089 study. The severity of CHE at entry to this cohort was mild or moderate in the majority of cases, and therefore differed from the entry criteria for BAP00089. It should be noted that only a minority of patients from BAP00089 were actually enrolled in BAP00091 (figures 6.3.2 and 6.3.3 of manufacturer's submission), so the potential for patient self-selection cannot be ruled out.

Of Cohort A, almost 80% of patients were successfully re-treated with 30 mg alitretinoin, as measured by PGA. However, it should be noted that there was a high proportion (69.2%) of patients who "responded and relapsed" to placebo treatment in the original trial also responded to "retreatment" with placebo in BAP00091 (though the absolute numbers of patients were small (9/13)). In addition, it should be noted the response rate as measured by PaGA was just 38%. The reason for such poor agreement between the PGA and PaGA evaluations for these patients (the 95% confidence intervals of the two estimates do not overlap) is unclear.

In Cohort B, comprising non-responders from BAP00089, 50.9%, 50.4% and 39.1% of patients who did not respond fully to initial treatment with placebo, 10 mg alitretinoin or 30 mg alitretinoin achieved clear/almost clear hands (p.49).

Table 6: Summary of Withdrawals from BAP00091

	Cohort B	Cohort A		
	Placebo	Placebo	30 mg	10 mg
Percentage of withdrawals	19.8% (n=48)	20.8% (n=14)	12.2% (n=6)	19.0% (n=4)
Main Reasons				
Insufficient response	5.3% (n=13)	17.0% (n=8)	4.1% (n=2)	-
Refused treatment/lack of co-operation	4.9% (n=12)	2.1% (n=1)	4.1% (n=2)	4.8% (n=1)
Adverse Events	4.5% (n=11)	4.3% (n=2)	4.1% (n=2)	4.8% (n=1)

As in BAP00089, in BAP00091 the proportion of withdrawals was substantial (Table 6), though the absolute numbers of patients involved was smaller.

Health-related quality of life data

Given the very limited detail provided on the collected DLQI data, the ERG asked for further details. The manufacturer's response indicated that 162 of the 319 patients (51.4%) entered into BAP00003 completed DLQI questionnaires both at baseline and 12 weeks. The manufacturer stated this reduced sample size meant that the trial was not sufficiently powered to show any statistically significant effect of alitretinoin treatment on DLQI, so therefore excluded this 'treatment effect' analysis from their submission. The main quality of life analysis presented in the manufacturer's submission of clinical efficacy examines the relationship between PGA state and DLQI, independent of treatment effect, using data from BAP00003 and an unpublished study of DLQI in [redacted] adults with CHE.[4] Both studies suggested some correlation between increasing mean DLQI and increasing severity of PGA state (see Tables 6.9.1 and 6.9.2), although mean DLQI scores for the matching PGA states differed somewhat between the studies, and the trend was less pronounced in the unpublished observational study than in BAP00003. Table 7 shows the reported similarities and differences between patients in the two studies, with BAP00003 including a much larger proportion of male patients with a shorter duration of disease (it should be noted that many other unknown variables might account for the different scores observed between these two separate studies). It should also be noted that around 65% of CHE patients in trial BAP00003 had PGA 'moderate' disease, whereas subsequent efficacy trials selected patients for inclusion on the basis of being in the 'severe' PGA state.

Table 7: Mean scores and patient characteristics for studies measuring DLQI

	BAP00003	Unpublished observational study (Augustin et al)[4]
n	162	[redacted]
Mean age (years)	48.2	[redacted]
% male	73.7	[redacted]
Duration of CHE (years)	3.1	[redacted]
PGA severe DLQI score	15.08	[redacted]
PGA moderate DLQI score	9.78	[redacted]
PGA mild DLQI score	5.93	[redacted]
PGA clear/almost clear DLQI score	1.74	[redacted]

The manufacturer stated that the minimally important difference (MID) DLQI threshold for patients with CHE was estimated to be 2.53, so concluded that improvement from severe to clear/almost clear PGA states indicates a meaningful improvement in quality of life. However, it should be noted that the manufacturer did not specify how the MID value of 2.53 was obtained.

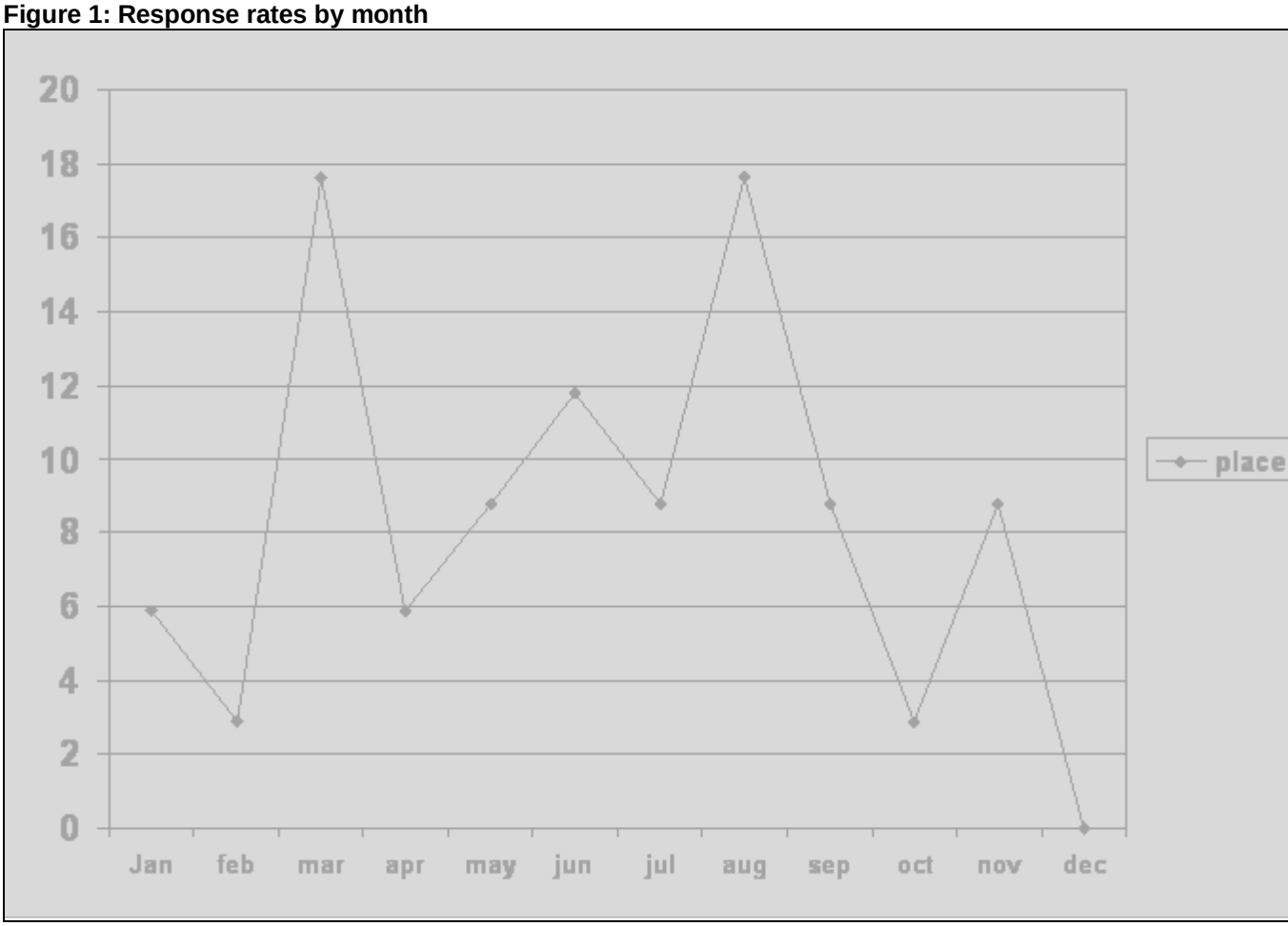
Other issues

The ERG requested further details on any potential underlying seasonal effects observed in the included studies. The manufacturer provided data on the proportion of PGA responses observed in placebo-treated patients in trial BAP00089 (Table 8).

Table 8: Summary by month of % PGA responses observed in placebo-treated patients in BAP00089 study

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Placebo	5.9	2.9	17.6	5.9	8.8	11.8	8.8	17.6	8.8	2.9	8.8	0

Figure 1: Response rates by month



It can be seen that placebo response rates are, in general, higher for the spring/summer months than for the winter months (Figure 1). However, because of the limited data available, it is not possible for the ERG to separate the influence of any potential seasonal effects from other factors, such as time on treatment.

4.2.2.2 Safety of Alitretinoin

As discussed in section 4.1.6., the adverse effects investigated and reported are appropriate. Analysis of safety relied heavily upon the phase III clinical study (BAP00089) which reported that 50% of the patient population experienced at least one adverse event (AE). Treatment-emergent AEs were more frequent in the 30 mg group than the 10 mg group, see Table 9. The most frequent adverse event was headache which was considered to be dose dependant (20% in 30 mg, 11% in 10 mg and 6% in placebo). For laboratory parameters there were increases in both total cholesterol and triglycerides, which occurred with greater frequency in the 30 mg dose than the 10 mg group – these may have implications for the risk of future

cardiovascular events. However, there is currently no long term evidence on the effect of alitretinoin on cardiovascular outcomes.

Table 9: Adverse event rates reported from trial BAP00089

	Alitretinoin 30 mg	**Alitretinoin 10 mg**	**Placebo**
	N=410	N=418	N=203
Any adverse event N (%)	244 (59.5%)	216 (51.7%)	101 (49.8%)
Serious adverse events N (%)	11 (2.7%)	17 (4.1%)	3 (1.5%)
Serious adverse events related to study treatment N (%)	4 (1%)	4 (15)	2 (1%)
Discontinuation due to adverse events N (%)	38 (9.3%)	22 (5.3%)	11 (5.4%)

Psychiatric status was measured with the validated and frequently used Centre for Epidemiologic Studies Depression scale (CES-D). Though the manufacturer's submission states that the General Health Questionnaire (GHQ) was also measured, data on this outcome were not available to the ERG. Details of a range of ophthalmologic outcomes were provided, including measures of visual disturbances, ocular motility and anterior segment changes. Details of x-ray evaluations of skeletal changes and bone mineral density measurements were also provided. Based upon the information supplied to the ERG there appears to be no evidence to suggest that alitretinoin has major safety issues in the short-term. The within study safety profiles are limited to a maximum of 48 weeks duration, thus there is no long-term data upon which to draw upon with respect to the safety profile of alitretinoin.

4.2.2.3 Efficacy and safety of comparator interventions

PUVA

Of the 13 RCTs identified, five were excluded (three did not separate between hands and feet, one was inadequately controlled and one did not adequately describe the results in English and probably had the wrong population)(See Appendix 2). The RCTs of PUVA in the treatment of CHE typically involved small numbers of patients and, with the exception of one trial comprising 158 patients were all in the range 12-44 patients. Of the eight RCTs considered for inclusion in the review a further four were excluded because they did not report patient-level data, thus four trials were included in a meta-analysis of response rates (clear/almost clear). The comparator for these studies was mostly, but not exclusively, UVB. Summary details of all eight trials (i.e. those included and excluded from the meta-analysis) are presented in Table 10 and Appendix 3.

The results of the meta-analysis gave a pooled odds ratio 0.72 (95% CI: 0.000005, 110990.51) (p.66). These excessively large 95% confidence intervals result from substantial clinical heterogeneity and significant trial heterogeneity and represent a result that is difficult to interpret in terms of determining any treatment effect of PUVA.

The qualitative synthesis conducted by the ERG (Table 10) found that none of the trials demonstrated superiority of PUVA over the comparator, but all trials reported improvements for the PUVA arm and in six of the eight trials this improvement form baseline was statistically significant. Whilst these findings indicate that PUVA may have some beneficial effects in the treatment of CHE, the very small sample sizes in these trials and lack of any placebo or no treatment comparison, they provide almost no information on the effects of PUVA and none that could be used in a formal indirect comparison.

In those trials of PUVA that reported relapse rates, the rates were 15% at 3 weeks, 50% at 10 weeks and 64% at 12 weeks. These were broadly in agreement with the clinician estimates of relapse of 10% at week 4, 20% at week 8, increasing to 40% at week 12 and 80% by week 20 (see Appendix 3).

Table 10: Summary of the eight included PUVA RCTs

Study	Comparison	Study population and severity scoring system	Beneficial effect of PUVA on symtoms?	Significant treatment difference?
Petering et al. 2003 N=27[5]	UVA vs. topical PUVA.	All patients had recurrent disabling bilateral symmetrical vesicular hand eczema for at least 3 months	Yes (change form baseline p<0.05)	No
Sezer et al. 2007 N=12[6]	Comparison of paint-PUVA on one hand and UVB on the other hand.	Subtype only CHE of dry and dyshidrotic types, (hyperkeratotic CHE excluded).	Yes (change form baseline p<0.05)	No
Rosen et al. 1987 N=35[7]	Oral PUVA (N=18) and UVB (N=17). One hand exposed the other an untreated control. .	Bilateral hand eczema, symmetrical distribution and severity of at least 6 months duration. Predominantly females (31/35) with vesicular CHE (26/31) enrolled. Two patients were hyperkeratotic in the PUVA arm.	Yes (change form baseline p<0.001)	Unclear
Simons et al. 1997 N=13[8]	UVB vs. topical PUVA on each hand.	Patients with vesicles or hyperkeratotic plaques of the hands present for > 6 months.	25% reduction in symptoms but unclear if significant	No
Sheehan-Dare et al. 1989 N=25[9]	UVA vs. PUVA	All patients had chronic eczematous changes on the palms for at least 6 months with either continuous or episodic vesiculation.	Yes and statistically significant but p value not reported).	Yes – in favour of UVA
Van Coevorden et al. 2004 N=158[10]	Oral PUVA at home N=78, Hospital administered bath PUVA N=80.	Chronic bilateral or unilateral hand eczema of at least 1 year's duration,	Yes but unclear if significant	No
Adams et al. 2007 N=15[11] (*Paper in German)*	One hand received topical PUVA and the other medium-dose UVA-1	Chronic dyshidrotic hand eczema.	Yes (p=0.0498).	Unclear
Grattan et al. 1991 N=15[12]	One hand received topical PUVA and the other UVA.	Recurrent disabling bilateral symmetrical vesicular hand eczema for at least 6 months	Yes (p<0.005)	

Ciclosporin

The one small (n=41) randomised study identified for ciclosporin found no difference in response rate between the ciclosporin and topical steroids over a 6 week period (summary details given in Appendix 4). The ciclosporin dosing regimen was relatively low (3 mg/kg/day) and the study population were not steroid refractory. In this study both ciclosporin and topical steroids improved eczema. The total disease activity score decreased to 57% of baseline (12.9 to 7.3) in the ciclosporin group (mean change -6, SD 4.3; $p<0.001$) and to 58% of baseline in the topical steroids group (mean change -5.7, SD 4.0; $p<0.001$), with no significant difference between treatment groups. Fifty percent of patients in both groups relapsed within 2 weeks (defined as an increase in disease severity score/extent of disease to >75% of baseline score).

Relapse rates in this study were of a greater magnitude, and the time to relapse was quicker than clinician estimates (after 2-week follow-up in the trial 50% of patients in the ciclosporin group had relapsed, compared with clinician estimates of 30% relapse by week 4, 50% by week 8 and 80% by week 12).

Comparing the results from this small trial with the estimates derived by the clinicians yields little. In this instance, the clinicians' opinion would appear to represent 'best evidence' for the effectiveness, or otherwise, for treatment using ciclosporin.

Azathioprine

No studies were identified that assessed the efficacy of azathioprine for CHE but, as for PUVA and ciclosporin, expert clinicians' opinions regarding likely values for efficacy, relapse and contraindications were elicited.

.

4.2.3 Summary

On the basis of the limited available evidence, alitretinoin appears to be a reasonably efficacious treatment for severe chronic hand eczema. Trial data indicate that around 43% to 53% of severe CHE patients (refractory to topical steroids) will respond to 12-24 weeks of treatment with 30mg alitretinoin, in terms of achieving a PGA state of clear/almost clear. Placebo response can clearly be observed among the included studies, though when compared directly against placebo the effect of alitretinoin remains statistically significant. However, the median time to relapse did not differ significantly between alitretinoin and placebo.

Retreatment with alitretinoin appeared to lead to remission in a substantial proportion of patients who had either not responded to the first course of treatment, or had responded but relapsed. However, the data suggest that a high level of response can be observed in remitting/relapsing patients who receive retreatment with placebo. In addition, there were unexplained inconsistencies between the physician and patient ratings of retreatment success.

In Cohort B, comprising non-responders from BAP00089, 50.9%, 50.4% and 39.1% of patients who did not respond fully to initial treatment with placebo, 10mg alitretinoin or 30mg alitretinoin achieved clear/almost clear hands (p.49).

The relevant comparators of interest are PUVA, ciclosporin, and azathioprine. There are no convincing evidence of the efficacy of these treatments on severe CHE, Furthermore, since alitretinoin has not been compared in a head-to-head trial against any of these comparators, and since none of the comparators have been compared against placebo in a severe CHE population, there is no direct or indirect clinical evidence on the effects of alitretinoin relative to PUVA, ciclosporin or azathioprine.

Dose-dependent headache was the most commonly reported adverse event in patients treated with alitretinoin. Serious adverse events were rare, but alitretinoin was associated with increases in both total cholesterol and triglycerides, which has implications for risks of future cardiovascular events. The main trial data is limited to a maximum of 48 weeks duration and therefore, there is no evidence currently available to assess the longer-term risks of alitretinoin treatment. Given that treatment with alitretinoin is likely to be intermittent and continual, long term safety data are required.

5 ECONOMIC EVALUATION

5.1 Overview of manufacturer's economic evaluation

The manufacturer's submission to NICE included:

1. A description of the systematic search of the economic literature conducted by the manufacturer (manufacturer's submission p.81, Section 10.3 and Appendix 3).

2. A report on the *de novo* economic evaluation conducted by the manufacturer (manufacturer's submission pp.81-118, Tables 7.2.1-7.3.5).

3. Base case analysis cost-effectiveness results from the model (manufacturer's submission pp.110-114, Tables 7.3.1-7.3.3).

4. One-way sensitivity analysis results from the model (manufacturer's submission pp.115-117, Tables 7.3.1-7.3.3).

5. Subgroup analysis results from the model (manufacturer's submission pp.114-115, Tables 7.3.4-7.3.5).

6. An Excel-based model comprising the manufacturer's electronic economic model.

Following requests from the ERG, the manufacturer provided the following:

1. Clarification on effectiveness data (including further details of the systematic review, PUVA trials, quality of life data, DLQI analysis, safety data, subgroup analysis, and other miscellaneous clarifications).

2. Clarification on cost-effectiveness data (including the utility mapping methodology, definition of relapse used in the model, assumptions used in the model, and other miscellaneous clarifications).

3. A revised Excel model with the inclusion of a 'placebo' arm, among other revisions.

4. Base case and subgroup analysis results from the revised model (manufacturer's response pp.13-17).

5.1.1 Natural history

Alitretinoin is indicated for use in the treatment of adults with severe CHE that is refractory to potent topical steroids (manufacturer's submission p.82). The manufacturer's model assumed that patients with severe CHE would begin treatment and start the model when their severe CHE was determined to be in the 'severe' PGA state. The model evaluated a heterogeneous cohort of patients aged 48 years old and weighing 81kg; 57% male and 15% of the total assumed to be women of child-bearing potential, which reflects the average characteristics of the patient group included in the BAP00089 trial (manufacturer's submission p.28).

A treatment course in alitretinoin was assumed to be given for between 12 and 24 weeks, depending on response (p.83). Patients responding to alitretinoin in each 4 week period were assumed to cease treatment immediately (including those responding after 4 or 8 weeks) Those patients who remained in the PGA 'severe' state after 12 weeks were assumed to withdraw from treatment and enter the refractory state of the model. Patients whose CHE was rated PGA 'clear' or 'almost clear' by 24 weeks were deemed to be in remission, whilst those whose CHE was rated PGA 'moderate', 'mild' or had returned to PGA 'severe' at 24 weeks were assumed to be refractory. Those in remission were assumed to relapse to a 'severe' PGA state after an average (median) time of 24 weeks (p.95). At this time the model assumed that a further treatment course in alitretinoin was given under the same assumptions as for the first course (p.83), although the transition probabilities between states were updated to reflect that patients were being retreated following relapse.

Those treated with alitretinoin were initially assumed to receive a dose of 30mg once daily. If an adverse event occurred (either headache or hyperlipidaemia) then it was assumed that for some patients the dose would be reduced to 10mg once daily, and for others treatment would continue uninterrupted. If a further adverse event occurred while on the lower dose then it was assumed that some patients would withdraw from treatment, entering the refractory state, while the remaining patients would again continue treatment uninterrupted. While not mentioned in the written submission, the model assumed only one adverse event could occur in each 4 week period.

The model's time horizon was assumed to be three years in the base case, although this was explored further in a sensitivity analysis (p.87).

For the comparator treatments, the treatment cycle was assumed to follow a similar pattern to that of alitretinoin. For ciclosporin, azathioprine and PUVA, the treatment course was assumed to be given for 16, 48 and 16 weeks respectively, with those patients with severe CHE still rated PGA severe withdrawing at 12, 16 and 16 weeks respectively (p.95). The average time to relapse for those responding to ciclosporin, azathioprine or PUVA was assumed to be 9.6, 10 or 18 weeks, respectively (p.95). However, it was assumed that patients experiencing an adverse event under any of the comparator treatments would not be offered a less-intensive second-line treatment (as under alitretinoin) but would immediately face the possibility of treatment being withdrawn – the adverse events considered for the comparators and the probabilities of withdrawal are given on pages 90-93 of the manufacturer's submission. Furthermore, for ciclosporin it was assumed that a maximum of four treatment cycles could be carried out on each patient, irrespective of the time horizon (p.95).

5.1.2 Treatment effectiveness within the submission

Data on response to treatment with alitretinoin were extracted from the phase III clinical trial BAP00089 for the first treatment cycle and from the follow-up cohort A of the phase III trial BAP00091 for subsequent treatment cycles (pp.93-94) (see earlier section 4.2 for a review of these trials). These data were collected at four week intervals and were modelled as such across each 24 week treatment period (manufacturer's submission p.93). Meanwhile, the data on disease progression for the comparators were based on "clinical expert opinion" (p.100).

Data on the number of adverse events and the probabilities of dose reduction or withdrawal from treatment were informed by either clinical trial BAP00089 (the proportion of patients reporting headache or raised cholesterol) or by the manufacturer's assumptions (probability of dose reduction and probability of withdrawal following adverse event, pp.90-92). Time to relapse following remission was informed by the BAP00089 clinical trial in the case of alitretinoin and by clinical opinion for the comparators (pp.96-97).

The manufacturer carried out two subgroup analyses as part of the submission. The first of these was in patients "in whom the CHE has predominantly hyperkeratotic features", whom "the SPC emphasises... are more likely to respond to alitretinoin treatment than those in whom the CHE predominantly presents as pompholyx"

(p.85). The manufacturer modelled this subgroup by "adjusting the efficacy data for alitretinoin to reflect the improved efficacy that has been observed in trials of predominantly hyperkeratotic patients treated with alitretinoin". Since four weekly trial data "were not available for the hyperkeratotic patient group", the manufacturer modelled the efficacy data for the first treatment cycle linearly over 24 weeks from "trials of predominantly hyperkeratotic patients treated with alitretinoin"– these trials were not explicitly identified. For subsequent treatment cycles with alitretinoin, the efficacy data "was based 4 weekly [sic] data for the overall population from BAP00091 because hyperkeratotic analysis was not available" (p.85). However, in response to a request for clarification by the ERG, the manufacturer confirmed that the response rate for subsequent cycles was in fact derived from a sub-group analysis of BAP00091 (p.7 of manufacturer's response).

The second sub-group analysis was in "women of child-bearing potential" (manufacturer's submission p.85). Due to the teratogenic effect of alitretinoin, such patients must abide by strict pregnancy-prevention rules (as specified in the SPC) and so receive additional pregnancy monitoring and contraception measures. While the efficacy of alitretinoin was assumed to be the same in these patients as in the base case, such patients were assumed to incur the additional costs associated with contraception and pregnancy consultation/testing (p.89).

The manufacturer did not carry out a subgroup analysis for patients at high risk of cardiovascular disease because it was felt that such patients would either "not be started on alitretinoin therapy" or "would start on 10mg and then titrate up to 30mg as per the SPC", and the manufacturer did not feel able to predict the relative proportion of patients managed in each of these ways nor "the rate at which elevated lipids would be brought under control at 10mg allowing the greater efficacy of 30mg to be modelled for subsequent treatment" (p.86).

5.1.3 Health related quality of life

Health-related quality of life weights (utility values) used in the base-case version of the model were derived in a two-stage process using data collected during the phase II trial BAP00003 (manufacturer's submission p.102) and a previously published algorithm examining the relationship between DLQI and EQ-5D in patients with psoriasis. Data from BAP0003 were reanalysed to examine the relationship between change in PGA state and DLQI, independent of treatment effect (manufacturer's

submission p.78 and Table 6.9.1). This was used to predict DLQI from the PGA scores observed in BAP00089. The manufacturer then employed an algorithm from Woolacott *et al.* (2006),[13] derived from data on psoriasis patients, to convert DLQI scores into EQ-5D utility scores (EQ-5D utility score = 0.956 – 0.0248 * DLQI score) (p.103).

The model applied the utility score associated with PGA state 'severe' to patients rated PGA severe and still receiving treatment and to those patients deemed to be refractory; the 'moderate' and 'mild' utility scores were applied to those patients receiving treatment rated moderate or mild on the PGA scale respectively; whilst the 'clear' and 'almost clear' utility scores were averaged to provide a single utility score which was applied to those patients in remission. Adverse events were assumed to have no impact on health-related quality of life. In the base-case analysis the model estimates quality-adjusted life years (QALYs) over a three year time horizon.

5.1.4 Resources and costs

The model considered the resource costs associated with patient treatment, monitoring and adverse events (manufacturer's submission p.104). These were identified from published sources (manufacturer's submission p.104).

Alitretinoin is priced at £411.43 per pack of 30 soft capsules (one capsule to be taken per day) for both 30mg and 10mg doses (manufacturer's submission p.83). In the model this was assumed to represent a cost of £383.88 every four weeks on the implicit assumption that the remaining two capsules were not wasted (manufacturer's submission p.88). Patients were assumed to cease treatment with alitretinoin as soon as they enter remission. Treatment with ciclosporin, PUVA and azathioprine was assumed to cost £164.64, £514.65 and £16.80 every four weeks, respectively (p.89).

The costs associated with monitoring of treatment and remission are given on pp.89-90 of the manufacturer's submission. Patients are assumed to visit the dermatologist once every four weeks whilst receiving any of the treatments under comparison. For patients receiving alitretinoin the monthly monitoring costs include lipid monitoring tests, and a pregnancy consultation and test for women of child-bearing age (assumed to represent 15% of the overall patient population). For ciclosporin the monthly monitoring costs include serum creatinine monitoring tests, for aziathoprone

liver function tests and TMPT monitoring, while for PUVA no additional tests are required. Remission is assumed to cost less following treatment with alitretinoin (£5.20 every four weeks) than following treatment with a comparator (£11.04 every four weeks) as patients are assumed not to receive topical steroids alongside emoilients.

The management costs associated with adverse events are given on pp.90-93 of the manufacturer's submission. In the base-case analysis the model estimates total costs to the NHS over a three year time horizon.

5.1.5 Discounting

The manufacturer's model applied a discount rate of 3.5% per annum to expected costs and health effects (p.106), in line with the NICE reference case.

5.1.6 Sensitivity analyses

The model did not employ probabilistic sensitivity analysis. The manufacturer's justification for this was that the clinical efficacy data for alitretinoin was "sourced from a single trial", whilst for the comparators it was sourced from "a single panel of clinical experts"; as such, the data "is highly uncertain and does not permit a meaningful characterisation of the uncertainty surrounding patient response to treatment" (p.108).

However, the manufacturer did explore (through one-way sensitivity analysis) the impact on the model's results of considering alternative time horizons (1 year; 6 years; 10 years; 20 years), a change in the efficacy of the treatments (a reduction in the efficacy of alitretinoin of 30%; an increase in efficacy of the comparator treatments of 50%) and alternative utility values. These utility values were estimated by applying the algorithm developed by Woolacott et. al (2006)[13] to DLQI values associated with each PGA score taken from an unpublished abstract for an observational study conducted in Germany[4] (manufacturer's submission p.79).

5.1.7 Model validation

The manufacturer's submission reports that the model was "double coded", with each coding reviewed by a person other than the person who constructed the model (p.109). Furthermore, the model was subjected to an "extreme value analysis" in which parameter values were varied beyond what would be considered "reasonable" and the simulated costs and utilities observed to ascertain if the model was consistent with the structural assumptions and *a priori* expected differences in costs and health benefits between the alternative treatments modelled (p.109).

The manufacturer's submission did not include a validation of the model results against the values observed in the clinical trials that informed the model, for example in terms of number of withdrawals from treatment. In addition, the manufacturer did not include an assessment of the stability of the results of the model, which was based on individual patient simulations, with respect to the number of patients simulated.

5.2 Critique of approach used

The manufacturer conducted a systematic search of the economic literature and identified no relevant prior studies on the cost-effectiveness of treatments for CHE. As such, the submission of a *de novo* economic evaluation was appropriate. The manufacturer's submission was built upon a Markov based patient-level simulation model constructed in Excel, consisting of a simple spreadsheet accompanied by a lengthy and relatively complex module of Visual Basic for Applications (VBA) code.

The ERG identified a number of shortcomings with the manufacturer's model.

5.2.1 Treatment effectiveness

One notable shortcoming with the model is that the efficacy data for those treatments other than alitretinoin were based on expert clinical opinion only. The submission provided a brief summary of this process (p.100), and the manufacturer provided more detail in response to a query from the ERG (manufacturer's response pp.17-19). In summary, the expert clinicians were presented with the publication of the BAP00089 trial, and were asked to estimate the distribution of patients between the

PGA states over six consecutive four week periods for each of the comparators. In addition they were asked about durability of response and time to relapse.

From the detail given by the manufacturer it is impossible to tell whether the clinicians were estimating a placebo-adjusted response or absolute rates of response. The clinicians were not asked to provide estimates of uncertainty in their estimates of efficacy. Consensus was achieved informally by round-table agreement, and so the weighting provided to each clinician's input is impossible to judge.

The reliability of the estimates of the efficacy of each of the comparator treatments generated by this elicitation process is subject to major doubts. The failure to adjust for the placebo response in the estimate of alitretinoin efficacy, and the lack of clarity about whether placebo response is included in the experts' estimates of efficacy for the comparators, means that the process may have biased the efficacy estimates in favour of alitretinoin. It is notable that when the results of the original model are combined with those of the revised model (which includes a placebo arm the efficacy of which was estimated using trial data rather than clinical opinion) then placebo appears to dominate azathioprine due to a higher rate of response with placebo (see results of additional analyses in chapter 6). The clinical plausibility of this is subject to doubt.

5.2.1.1 Relative efficacy

The model originally submitted to NICE did not include a "supportive care" (or "placebo") arm and the treatment effects for alitretinoin were not placebo adjusted; as such, the model did not address whether alitretinoin was a cost-effective alternative to supportive care. Whilst a supportive care arm was not specified as a requirement in the scope, discussions with the ERG's clinical advisor suggested that supportive care (including topical corticosteroids even with sub-optimal response) may be considered a relevant alternative, particularly in patients no longer eligible for immunosuppressants or PUVA.

In response to a request from the ERG, the manufacturer provided a revised model with a "placebo" arm (i.e. effectively a 'do nothing' option). The efficacy data for this arm were taken from BAP00089 and BAP00091. For the first treatment cycle the four-weekly data from BAP00089 were used directly, whilst for subsequent cycles the

efficacy data were derived from a separate analysis "performed to understand the efficacy of patients responding to placebo in BAP00089 and then receiving placebo in BAP00091". These data represented 13 patients and were available only for the 24-week time point; as such, "the data were linearly allocated over a 4 weekly time period" (p.13 of manufacturer's response).

The manufacturer incorporated further changes into the revised model they developed following the ERG's points of clarification. At the request of the ERG, the alitretinoin arm was assumed to include the cost of two blood tests from TSH monitoring over the course of each treatment cycle (£3 each). However, the revised model was restricted to comparing only alitretinoin with placebo (in the Excel model the azathioprine arm had been overwritten with the placebo arm and the other comparator arms had been effectively disabled and hidden from view) and adverse events were removed from the model without explanation (this was noted on p.13 of the manufacturer's response). The removal of adverse events from this placebo-adjusted model means that the costs of these events are omitted and furthermore no patients will discontinue treatment or move to the lower dose of alitreinoin as a result.

5.2.1.2 Relapse

The assumption given in the manufacturer's written submission that all patients will "re-enter the severe state" upon relapse (manufacturer's submission p.96) does not appear to be correctly implemented in the model's VBA code (section 5.2.5). It seems the first 4 weeks of each subsequent treatment cycle are omitted, with patients 'relapsing' to the PGA state in which they would be after 4 weeks of treatment. However, it is not clear that such an assumption is valid in any case. The relevant clinical trials defined relapse as a "return to 75% of baseline mTLSS" – the disease activity score used in the BAP00089 and BAP00091 clinical trials. Following a request by the ERG, the manufacturer confirmed that 30.6% of those patients in the 30mg alitretinoin group who had "just relapsed by attaining 75% of their baseline mTLSS in [trial] 089" were PGA moderate, with the remaining 69.4% PGA severe (p.25 of manufacturer's response). Thus the clinical trial data used to inform response to second line treatment is derived from a less severe patient population than that modelled, and may overestimate the response rates in patients who do restart treatment once their CHE is rated as PGA severe (see earlier section on 'Response to retreatment' section 4.2.2 for further detail on BAP00091).

Consultation with a clinical expert suggested that in clinical practice patients may be initiated on subsequent treatment cycles with less severity of CHE than that required for the initial course because a successful response has already been demonstrated. As such, it appears reasonable that the model should reflect the definition of relapse used in the clinical trial. The ERG modified the model's VBA code so that following relapse patients entered the severe and moderate states in the proportions observed in BAP00089 – this analysis is discussed in chapter 6.

5.2.1.3 Adverse events and withdrawal

The model considered the two most common adverse events with alitretinoin, headache and hyperlipidaemia, but did not include any others (Table 7.2.5, p.90 manufacturer's submission). Hyperlipidaemia encompasses both raised cholesterol and raised triglycerides, but the probability of hyperlipidaemia was based on the number of patients reporting raised cholesterol and did not include hypertriglyceridaemia, which was reported in 8% of patients given alitretinoin 30mg in BAP00089. Headache was assumed to be managed with paracetamol. Hyperlipidaemia was assumed to be managed with statins for four weeks and two GP visits, and the long-term consequences of any untreated hyperlipidaemia were not considered. The impact of hyperlipidaemia on withdrawal and treatment costs may be underestimated due to the exclusion of the increased risk of raised triglycerides (see pp.68-69 and Table 6.7.3 of manufacturer's submission for more detail on rates of adverse events).

The first cycle of alitretinoin is assumed to be associated with a 20% risk of headache and 14% risk of hyperlipidaemia. Upon experiencing a first adverse event, 20% of those with headache and 40% of those with hyperlipidaemia are assumed to switch to a lower dose of alitretinoin. This is a modelling assumption and does not reflect the practice in the clinical trials that inform the model. (p.80 of manufacturer's submission). However, the SPC for alitretinoin suggests that in practice dose reduction could be used to manage adverse events, and this was confirmed by consultation with a clinical expert. Those patients that switch to the lower dose of alitretinoin (10mg) are then assumed to face a lower risk of adverse events (11% for headache and 3% for hyperlipidaemia). Those patients on the lower dose that

experience a subsequent adverse have a 20% probability of withdrawal due to headache and a 40% probability of withdrawal due to hyperlipidaemia.

As a result of a potential modelling error, patients face only approximately 9% chance of experiencing a headache over the 24 week treatment cycle and 6% chance of hyperlipidaemia with 30mg alitreinoin. The ERG attempted to evaluate the number of patients withdrawing from treatment in the first treatment cycle as a result of these assumptions; it would appear that in the first treatment cycle only 4.4% of patients would switch to the lower dose, and only 0.01% would actually withdraw from treatment. This contrasts with the number of withdrawals observed in the clinical trial (26%), although only 9.5% of patients withdrew due to an adverse event (see table 2, section 4.2.2). In addition, the assumptions imply 19 adverse events per 100 patients in the first treatment cycle (allowing for repeated events on the same patient). This also contrasts with the clinical trial results in which 60% of patients reported any adverse event while receiving 30mg alitretinoin.

The adverse events considered for ciclosporin, PUVA and azathioprine were based on the manufacturer's assumption informed by SPC and published data where available. The management of hyperlipidaemia is assumed to be more expensive for patients receiving ciclosporin in comparison to those receiving alitretinoin (£20.07 and £14.40), although the reasons for this are unclear.

5.2.2 Health-related quality of life

As noted in the submission (p.102), the DLQI is a widely used measure of quality of life in patients with dermatological diseases. Unfortunately no DLQI data were collected during the phase III study, so the model relied on data from the phase II dose ranging study (which did not include the 30mg dose of interest).

The manufacturer stated that the phase II study was not sufficiently powered to demonstrate statistically significant differences in DLQI between the 10mg, 20mg, 40mg and placebo groups at the end of treatment; however, the manufacturer carried out an analysis of the change in DLQI associated with change in PGA status (see earlier 'Health-related quality of life data', section 4.2.2). Data from a subset of 162 patients included in the BAP0003 trial were available and were analysed using a generalized mixed model, with treatment group and PGA score at 3 months included as fixed effects and investigational centre included as a random effect

(manufacturer's submission p.102). Change in DLQI associated with change in PGA status was found to be highly statistically significant (manufacturer's submission, Tables 6.9.1 and 7.2.14). The estimated relationship was then used to predict the DLQI score on the basis of the PGA scores recorded for patients in the BAP00089 trial. These predicted DLQI scores were then converted to EQ-5D scores using the previously published algorithm developed using data from patients with psoriasis.[13] The use of this two-stage prediction method, which incorporates an algorithm developed on a set of patients with psoriasis, means that the utility values used in the submission should be interpreted with caution. The uncertainty in the two-stage prediction is not reflected in the model outputs. The manufacturer refers to data from an unpublished abstract that directly links DLQI with PGA state:[4] these were used by the manufacturer in a sensitivity analysis and could be viewed as a more appropriate basis for predicting EQ-5D scores (see earlier 'Health-related quality of life data', section 4.2.2 for more detail).

The model conflates the "clear" and "almost clear" PGA states and assumes the same number of patients in each. The manufacturer confirmed that fewer than 50% of patients who were in these two PGA states at the end of alitretinoin treatment were in the clear state (46% for the 30mg group and 34% for the 10mg group) – this compares to 17% for the placebo group (p.20 of manufacturer's response). Since conflating these two groups artificially favours treatments with relatively more patients in the almost clear group, this assumption results in a bias *against* alitretinoin when compared to placebo in the manufacturer's revised model, resulting in an increase in the ICER.

5.2.3 Resource use and costs

The submission is clear that a treatment course for alitretinoin may be given for 12 or 24 weeks, depending on response (manufacturer's submission p.83). As such, one would expect that patients entering remission after 4 or 8 weeks would continue treatment (and continue incurring the associated costs) until at least 12 weeks into each treatment course – this has been confirmed by the clinical advisor to the ERG, who suggested that treatment might continue in such patients due to concerns about relapse upon discontinuation. However, the base-case model assumes that alitretinoin patients entering remission immediately stop the treatment, thus ceasing to incur the costs of treatment. This assumption was not discussed in the manufacturer's written submission (although it was acknowledged on p.29 of the

manufacturer's response) and would underestimate the cost of prescriptions, resulting in a bias in favour of alitretinoin.

When this assumption is combined with fact that in the model patients may enter remission as soon as they relapse, a proportion of patients re-treated with alitretinoin never incur the cost of the drug. The ERG calculated that including prescription and monitoring costs for patients in remission at weeks 4 and 8 of each treatment cycle would increase the cost of alitretinoin by approximately £528 per patient over the course of three years. This is in addition to the underestimation of prescription costs caused by omitting the first 4 weeks of every treatment cycle.

Topical steroids are included in the remission costs for all treatment comparators except alitretinoin. After consulting a clinical expert, the ERG believes that differences in management for patients not receiving active treatment would not exist in clinical practice. Therefore this assumption may underestimate the true costs of managing patients with repeated cycles of alitretinoin, biasing the results in favour of alitretinoin.

5.2.4 Sub-groups

The sub-group analysis undertaken in patients with hyperkeratotic hand eczema was not accompanied by an analysis examining the remaining sub-group with pompholyx CHE. Furthermore, it again employed the absolute rate of response to alitretinoin as the estimate of treatment efficacy without adjustment for placebo response. The review of the clinical effectiveness evidence did not uncover firm evidence that any sub-group would perform better than the overall patient population (see Table 5, section 4.2.2). At the request of the ERG the manufacturer provided the sub-group analysis with alitretinoin compared to a supportive care arm and conducted an additional analysis in patients whose CHE was characterised as hyperkeratotic and pompholyx. The manufacturer did not provide a sub-group analysis in patients characterised as pompholyx only, stating that too few patients of this category were included in the clinical trial. Response rates for the sub-group analyses were calculated from the relevant patients included in the BAP00089 and BAP00091 trials in the same manner as that used for the overall patient population.

A recent paper by Diepgen et al. (2009), based on clinical data for 416 patients with hand eczema from 10 European patch test clinics, found a female preponderance

was most pronounced in the young age groups and the proportion of women less than 40 yrs of age was 34% of the total.[14] Whilst it is acknowledged that these patients had 'hand eczema', rather than 'chronic hand eczema refractory to topical steroids', it suggests that a statistic of 15% may be an underestimate for clinical practice. The written submission reports that "one pregnancy occurred during clinical trials with alitretinoin in a patient who failed to comply with the defined contraceptive measures. The pregnancy was terminated and failure of the pregnancy prevention program was reported as an SAE" (manufacturer's submission p.70). It should be noted that the model does not consider the effects on NHS/PSS costs or on patients' health-related quality of life of a termination being required in the case of pregnancy. Whilst the impact on health-related quality of life is difficult to establish, the NHS reference costs for 2006/7 suggest that a termination of pregnancy in an elective inpatient setting costs an average of £660 (HRG code MA182).

5.2.5 Other issues

5.2.5.1 Model validation/execution

A major shortcoming of the model is the heavy reliance upon poorly annotated VBA code, which significantly diminishes its transparency and verifiability – it was difficult to verify that the model behaves as described, and some important assumptions appear to be hard-coded into the VBA code and were either not reported in the written submission or were not implemented as described.

The written submission states that, after a patient relapses, "…it has been assumed for this model that they will at that point re-enter the severe state" (manufacturer's submission p.96). However, analysis of the VBA code and the model's output suggests that this is not the case and that most patients will resume treatment in non-severe states (with some patients immediately re-entering remission). This code is discussed further in Appendix 5. A brief inspection of the model's base-case output for alitretinoin reveals that, whilst in week 0 (the start-date of the first treatment cycle) all patients are assumed to be in the severe state, in weeks 36, 68, 100, and 132 (the start-dates for subsequent treatment cycles for those patients who relapse) patients continuing with treatment are distributed among the severe, moderate, mild and remission states, with those patients 'relapsing' to non-severe states not incurring the utility associated with the severe state at any point during the transition. Analysis of the VBA code reveals that 4 weeks has been subtracted from the time to remission,

with the effect that: (a) patients immediately resume treatment in the state that they should in fact enter four weeks into the retreatment cycle, and (b) each subsequent treatment cycle is four weeks shorter than the initial treatment cycle. In summary, week 0 of every re-treatment cycle has been omitted, with patients immediately entering week 4. The ERG attempted to amend the VBA code to fix this issue (see chapter 6).

Analysis of the model's VBA code suggests that the time to relapse for all patients measured from 12 weeks into the treatment course for *all* patients (whether they entered remission before or after this time). While the decision to measure time to relapse from 12 weeks for all patients *may* be justifiable on the basis that patients are assumed to enter relapse throughout the 24 week treatment course and the time to relapse is a median figure, assuming the same date of relapse for all patients would appear to negate one of the key benefits of conducting a patient-level simulation.

The model takes the form of a patient level simulation that by default evaluates 10,000 patients; this is 100 times more than the cohort size entered in the spreadsheet since the VBA code is hard-coded to perform 100 replications for each patient. These assumptions are not mentioned in the written submission and no justification is provided for the numbers selected. The manufacturer did not present pseudo-standard errors that one would expect to accompany a patient-level simulation in order to describe the size of the simulation error. As such, the amount by which the results could differ if the simulation were repeated with a different set of random numbers is unclear. The manufacturer's revised model used the random seed 2 within the VBA code, and generated an ICER for alitretinoin of £12,931 per QALY. A set of five simulations undertaken by the ERG using the arbitrarily selected random seeds 10, 20, 333, 500 and 999 produced ICERs for alitretinoin of £13,031, £12,718, £12,989, £12,931 and £12,697 per QALY, respectively. While this has not revealed any significant changes in the ICER (average difference £120, maximum difference £234), it by no means represents a full investigation of the simulation uncertainty. In order to understand the amount by which the ICER could vary due to simulation error, it would be necessary to compute pseudo standard errors.

The Excel model assumes that all ciclosporin patients move to the refractory state after 80 weeks, even though some are in remission at 76 weeks. This assumption is not discussed in the written submission and will underestimate the QALYs associated with ciclosporin, which results in bias in alitretinoin's favour.

5.2.5.2 Decision uncertainty

The manufacturer provides a limited series of one-way sensitivity analyses that cannot fully characterise the uncertainty in the estimates of costs and effects. A major source of uncertainty is likely to be the health-related quality of life estimates included in the model. Another source of uncertainty is in the reliability, interpretation and accuracy of the efficacy values estimated for the comparators.

5.2.5.3 Other issues

Inspection of the base-case output suggests that a half-cycle correction may be warranted – this was rejected in the written submission due to the "uncertainty surrounding patients' transitions through the model" (manufacturer's submission p.97). Such a correction is likely to modestly reduce the ICER for alitretinoin since the transitions to states with higher utility are more rapid than with the treatment comparators.

The results given in the manufacturer's written submission and calculated by the manufacturer's model are not fully incremental, consisting of pair-wise comparisons of alitretinoin versus each comparator treatment in turn. A fully incremental analysis was carried out by the ERG (chapter 6).

5.3 Results included in manufacturer's submission

In the manufacturer's submission, the base case ICERs reported for alitretinoin were £8614 per QALY versus ciclosporin, -£469 per QALY versus PUVA (with alitretinoin dominant) and £10,612 per QALY versus azathioprine.

Where the time horizon was shortened to 1 year, these ICERs increased (to £17,756 per QALY for alitretinoin versus azathioprine), whilst over longer time horizons these ICERs generally fell (to £9,324 per QALY for alitretinoin versus azathioprine over a 10 year time horizon). While the reported ICER for alitretinoin versus ciclosporin is slightly higher for the 20 year horizon than for the 10 year horizon, this would appear to be due to random variation.

Where the utility values used in the model were replaced with those derived from the Augustin et al. study,[4] the ICERs rose significantly (to £22,312 per QALY for alitretinoin versus azathioprine).

In patients with hyperkeratotic CHE, the ICERs for alitretinoin versus the comparators were higher than in the base case (£11,177, £-183 and £13,174 per QALY versus ciclosporin, PUVA and azathioprine respectively) due to the alternative efficacy data adopted. In women of child-bearing potential, the ICERs for alitretinoin versus the comparators were also higher than in the base case (£9,109, £54.27 and £11,038 per QALY versus ciclosporin, PUVA and azathioprine respectively) due to additional costs associated with pregnancy prevention for patients receiving alitretinoin.

5.3.1 Results of the revised model

Following requests from the ERG, the manufacturer produced a revised model which responded to some of the ERG points of clarification. In particular, it included a supportive care (or "placebo") arm; however, the remaining treatment comparators were removed from the model, as were adverse events. The results of this revised model are given in the manufacturer's response pp.13-17 and summarised below and in Table 11.

Base case

As described above, the main feature of the manufacturer's revised model was the comparison of alitretinoin with a placebo ('do nothing') option. In the base case scenario, with a 3 year time horizon, alitretinoin has an ICER of £12,931 per QALY gained against placebo. Where the time horizon is shortened to 1 year, the ICER rises significantly to £21,562 per QALY; conversely, as the time horizon increases, the ICER falls, albeit at a diminishing rate, reaching £10,765 per QALY for a time horizon of 20 years.

These results may be explained in the following way. For longer time horizons, the utility benefits resulting from response to treatment (including, crucially, the potential utility benefits resulting from the possibility of successful retreatment following relapse) are more fully captured: since alitretinoin has a better response rate than placebo over each treatment and retreatment cycle this results in a lower ICER for alitretinoin as the time horizon is extended. However, over time the number of

patients considered refractory and therefore not considered for further treatment cycles increases, and so the *absolute* differences in incremental cost and utility benefit between alitretinoin and placebo in each treatment cycle become less pronounced, and this effect is further exacerbated through discounting – as a result, the impact on the ICER of extending the time horizon becomes smaller over time.

An increase in the withdrawal rate for alitretinoin would result in the ICER plateauing more quickly and at a higher value than that demonstrated here. As the manufacturer appears to have underestimated withdrawal due to alitretinoin adverse events in the original model, and entirely removed the possibility of such withdrawal in the revised model, the ICERs reported here should be treated with some caution – the incorporation of a withdrawal rate due to alitretinoin adverse events commensurate with that observed in the clinical trials would likely increase the ICER for alitretinoin versus placebo, although it is not known to what extent.

Table 11: Results of manufacturer's revised economic model

Scenarios	Treatment	Total Costs	Incremental Costs	Total QALYs	Incremental QALYs	ICER
Base Case	**Placebo**	£611.83		1.79		
	Alitretinoin	£3,391.98	£2,780.15	2.01	0.22	£12,930.96
1 year	**Placebo**	£313.55		0.65		
	Alitretinoin	£2,207.96	£1,894.41	0.74	0.09	£21,562.06
6 years	**Placebo**	£995.00		3.32		
	Alitretinoin	£4,432.32	£3,437.32	3.63	0.31	£11,171.56
10 years	**Placebo**	£1,438.95		5.12		
	Alitretinoin	£4,975.34	£3,536.39	5.44	0.32	£10,967.78
20 years	**Placebo**	£2,315.14		8.67		
	Alitretinoin	£5,969.17	£3,594.03	9.01	0.34	£10,765.49

QALY = quality-adjusted life-years; ICER = incremental cost-effectiveness ratio

Health related quality of life

The manufacturer provided a further sensitivity analysis in which they assumed that utility value for patients in remission would be that associated with the 'almost clear' PGA state (0.88) rather than an average of the 'almost clear' and 'clear' states. This increased the ICER for alitretinoin compared to placebo to £14,024.85. When TSH monitoring was added the ICER rose modestly to £14,060.52.

Sub-group analyses

The manufacturer's revised estimate of the cost-effectiveness of alitretinoin compared to placebo in sub-group analyses is shown in Table 12 below. In hyperkeratotic patients, no sensitivity analysis was performed on the time horizon. In the single analysis performed (with a 3 year time horizon), alitretinoin had an ICER of £15,019 versus placebo – this is higher than in the base case due to the alternative efficacy data used.

The manufacturer presented the comparison of alitretinoin with all of the relevant comparators for the sub-group of patients with CHE characterised as hyperkeratotic and pompholyx (p.7 of manufacturer's response). As sub-group data were not available for the comparators (PUVA, ciclosporin and azathioprine) the efficacy was unchanged from the original model.

Table 12: Results of the manufacturer's revised economic model: sub-group analysis

Scenarios	Treatment	Total Costs	Incremental Costs	Total QALYs	Incremental QALYs	ICER
Hyperkeratotic	**Placebo**	£585.44		1.76		
	Alitretinoin	£3,419.91	£2,834.47	1.95	0.19	£15,018.95
Hyperkeratitic and pompholyx	**Placebo**	£566.81		1.76		
	Alitretinoin	£2,867.43	£2,300.62	1.84	0.08	£26,013.22

5.4 QALY = quality-adjusted life-years; ICER = incremental cost-effectiveness ratio

Comment on validity of results presented with reference to methodology used

As discussed in section 5.2.1, a major shortcoming with the model is that the efficacy data for those treatments other than alitretinoin were based on expert clinical opinion only. The use of expert opinion may be justified where trial data do not exist to inform the relevant parameters, but it should be elicited in a methodologically rigorous manner. Whilst the manufacturer provided more detail of this process in response to a query from the ERG (pp.17-19 of manufacturer's response), the ERG remains unconvinced that this elicitation process generated reliable estimates of the efficacy of each of the comparator treatments, and may well have underestimated each comparator treatment's efficacy (to the apparent benefit of alitretinoin). When the base-case results of the original model are combined with those of the revised model (which includes a placebo arm whose efficacy was estimated using trial data rather than clinical opinion) then placebo appears to dominate azathioprine (see chapter 6).

As such, the ERG does not regard the ICERs generated by the original model as providing a reliable indication of the cost-effectiveness of alitretinoin compared to each of the comparators considered. The comparison of alitretinoin with placebo made in the revised model is of greater merit given the more reliable efficacy data in the comparator arm. However, the omission of adverse events entirely from this model, in combination with a number of other factors, means that it underestimates the costs of treatment with alitretinoin and so the true ICER may be higher.

Serious issues remain around the implementation of the model in Excel. The manufacturer has made substantial use of sparsely annotated VBA code and does not appear to have implemented a number of the assumptions given in the written submission correctly (see chapter 5, section 5.2.5). The ERG has attempted to amend the VBA code in places to provide more appropriate estimates of the ICERs (see chapter 6) but in some cases this was not feasible. In addition, it is unlikely that the issues discussed in chapter 5 represent an exhaustive summary of the problems with the model due to the unnecessary complexity and opacity of the VBA code used.

5.5 Summary of uncertainties and issues

Health related quality of life

As the relief of symptoms and consequent improvement in health related quality of life is the aim of treatment for chronic hand eczema, the ERG believes that the economic evaluation of alitretinoin should be based on good evidence of the improvement in health related quality of life offered by alitretinoin. However, the estimates used in the submission are subject to a great deal of uncertainty due to the two-stage prediction employed and the paucity of direct observations in the population of interest.

Length of treatment with alitretinoin

It will cost the NHS £411.43 per patient for one month of treatment with alitretinoin and so it is important that the economic evaluation reflects the true cost of providing alitretinoin to patients with chronic hand eczema. The manufacturer assumes that patients receiving alitretinoin visit the dermatologist every four weeks and cease treatment as soon as they respond, even if this is after only four or eight weeks of treatment. If in practice patients would receive treatment for longer then the manufacturer will have significantly underestimated the costs to the NHS.

Patient population

The manufacturer assumes that patients with severe chronic hand eczema would be offered treatment with alitretinoin only when their disease is rated as severe on the PGA score, both for initial treatment and treatment for relapse. It is unclear to the ERG whether this reflects the population of patients with steroid refractory chronic hand eczema to which clinicians would aim to provide treatment.

Execution of the decision analytic model

The length of follow-up in the clinical trials was up to 48 weeks. The manufacturer employed a decision analytic model to extrapolate the effect of alitretinoin over multiple treatment courses up to a time horizon of 3 years, and to compare alitretinoin with other relevant treatments including PUVA, ciclosporin and aziathoprine. Inspection of the decision analytic model has revealed errors in execution that suggest that it does not reflect the assumptions made by the manufacturer in the written submission. In particular, the first four weeks of every treatment cycle bar the initial cycle are omitted. Furthermore the definition of relapse in the model does not correspond to that used in the relevant clinical trials. As a consequence the estimated costs and health outcomes presented by the manufacturer may be regarded as unreliable.

Adverse events and withdrawal

The economic evaluation originally presented by the manufacturer incorporates a lower rate of adverse events than that observed in the clinical trials and a lower rate of withdrawal from treatment resulting from adverse events, overestimating the number of patients who could potentially benefit from alitretinoin. This was exacerbated by the removal of adverse events from the revised model

Comparators

The ERG regards the comparisons of alitretinoin against azathioprine, ciclosporin and PUVA made in the original submission to be of limited value given that the efficacy data for those comparators were based on expert clinical opinion only. The ERG remains unconvinced that the elicitation process generated reliable estimates of the efficacy of each of the comparator treatments. As such, the ERG does not regard the ICERs generated by the original model as providing a reliable indication of the cost-effectiveness of alitretinoin compared to each of the comparators considered. The comparison of alitretinoin with placebo made in the revised model is of greater merit given the more reliable efficacy data in the comparator arm.

6 Additional work undertaken by the ERG

Fully incremental analysis

The results given by the manufacturer were not fully incremental, consisting of pair-wise comparisons between alitretinoin and each of the other treatment comparators.

Table 13 provides the results of a fully incremental base-case analysis conducted by the ERG using the originally submitted model.

Table 13: Incremental cost-effectiveness analysis of manufacturer's original results

Treatment	Cost (£)	QALYs	ICER (£ per QALY)
Azathioprine	805.25	1.75	N/A
Ciclosporin	1,580.72	1.79	ED by alitretinoin
PUVA	3,481.28	1.80	D by alitretinoin
Alitretinoin (30mg)	3,388.33	2.00	10,612 (vs azathioprine)

QALY = quality-adjusted life-years; ICER = incremental cost-effectiveness ratio; ED = ruled out by extended dominance; D = dominated; N/A = not applicable

Integrating the supportive care arm given in the revised model into a fully incremental analysis is straightforward since the manufacturer removed adverse events from the revised model and did not report on the adverse event profile associated with supportive care. Removing adverse events from the original model allows a fully incremental analysis to be carried out with the inclusion of the supportive care arm from the revised model, and the results of this analysis are provided in Table 14.

Table 14: Incremental cost-effectiveness analysis of manufacturer's original results combined with placebo

Treatment	Cost (£)	QALYs	ICER (£ per QALY)
Azathioprine	852.08	1.76	D by supportive care
Supportive care	611.83	1.79	N/A
Ciclosporin	1,690.83	1.80	ED by alitretinoin
PUVA	3,641.94	1.80	D by alitretinoin
Alitretinoin (30mg)	3,391.98	2.01	12,931 (vs supportive care)

QALY = quality-adjusted life-years; ED = ruled out by extended dominance; D = dominated; N/A = not applicable

Additional analyses

A number of additional analyses were carried out by the ERG on the revised model (comparing alitretinoin with supportive care). These were performed separately for two sets of utility data: those derived from the BAP0003 phase II trial population, used by the manufacturer in the base-case analysis; and those derived from an unpublished abstract of a German observational study (Augustin et al)[4] that directly links DLQI with PGA state, used by the manufacturer in a sensitivity analysis on the original model. In both cases the DLQI values were converted to EQ-5D utility weights using the algorithm from Woolacott *et al.* (2006)[13] reproduced on p.103 of the manufacturer's submission. These values are reproduced in Table 15 below:

Table 15: Comparison of utility estimates estimated by PGA score

PGA State	BAP0003		Augustin	
	DLQI	Utility	DLQI	Utility
Severe	15.08	0.582	[redacted]	[redacted]
Moderate	9.78	0.713		
Mild	5.93	0.809		
Clear / almost clear	1.74	0.913		

Furthermore, following advice from the ERG's clinical adviser, it was assumed that all patients *except* potentially child-bearing women would see a dermatologist once every 6 weeks with alitretinoin treatment and every 12 weeks under supportive care (rather than once per month as in the manufacturer's model). Potentially child-bearing women were assumed to see a dermatologist once every 4 weeks if they were receiving alitretinoin treatment, or once every 12 weeks otherwise. In addition, again following advice from the ERG's clinical advisor, patients in remission after receiving alitretinoin treatment were assumed to incur the cost of topical steroids, in line with alitretinoin's treatment comparators.

The ERG undertook four additional analyses, the results of which are shown in Table 16, which follows a brief description of each of the additional analyses:

1 Base case reanalysis using alternative utility weights and less frequent dermatologist visits

A reanalysis of the base case was carried out to examine the sensitivity of the results of the revised model to changes in the utility values associated with each of the PGA states and the impact of less frequent dermatologist visits.

2 Patients relapse into PGA moderate and severe

As discussed in section 5, the written submission states that, after a patient relapses, "…it has been assumed for this model that they will at that point re-enter the severe state" (manufacturer's submission p.96). However, analysis of the VBA code and the

model's output suggests that this is not the case and that most patients will resume treatment in non-severe states, with some patients immediately re-entering remission.

As reported in section 5, it is not clear that such an assumption is valid – the clinical trials utilised a definition of relapse as a "return to 75% of baseline mTLSS". Following a request by the ERG the manufacturer confirmed that 30.6% of those patients in the 30mg alitretinoin group in BAP00089 who relapsed were PGA moderate, with the remaining 69.4% PGA severe (p.25 of manufacturer's response).

The ERG modified the VBA code so that patients relapsed to the appropriate PGA state (30.6% of relapsing patients into the moderate state and the remainder into the severe state - see Appendix 6 for more details).

3 (a) Potentially child-bearing women only; and (b) Men only

The ERG performed two further analyses considering (a) potentially child-bearing women only, and (b) men only. These assumptions were fed into the model on the 'Inputs' worksheet given in the spreadsheet by setting the 'child bearing women' cell and the 'sex distribution (% male)' cell to either 0% or 100%, as required. The men only analysis can be generalised to post-menopausal women.

4 Reinstate adverse events for alitretinoin only

The manufacturer removed adverse events from the revised model without justification (see section 5). In the absence of any adverse event profile for supportive care, the ERG performed an analysis on this revised model with adverse events reinstated from the original model for alitretinoin only – this provides a 'worst case' scenario for alitretinoin in this regard.

Table 16: Results of additional analyses

	BAP0003 utility data			Augustin utility data		
Analysis 1: Base-case reanalysis						
Treatment	**Cost**	**QALYs**	**ICER (per QALY)**	**Cost**	**QALYs**	**ICER (per QALY)**
Supportive care	£481.40	1.79		£481.40	2.05	
Alitretinoin (30mg)	£3,369.21	2.01	£13,431.67	£3,369.21	2.16	£27,996.89
Analysis 2: Patients relapse into PGA moderate and severe						
Treatment	**Cost**	**QALYs**	**ICER (per QALY)**	**Cost**	**QALYs**	**ICER (per QALY)**
Supportive care	£481.60	1.78		£481.60	2.05	
Alitretinoin (30mg)	£3,509.33	1.99	£14,525.65	£3,509.33	2.15	£29,864.39
Analysis 3a: Potentially child-bearing women only						
Treatment	**Cost**	**QALYs**	**ICER (per QALY)**	**Cost**	**QALYs**	**ICER (per QALY)**
Supportive care	£481.40	1.79		£481.40	2.05	
Alitretinoin (30mg)	£3,548.95	2.01	£14,267.64	£3,548.95	2.16	£29,739.38
Analysis 3b: Men only						
Treatment	**Cost**	**QALYs**	**ICER (per QALY)**	**Cost**	**QALYs**	**ICER (per QALY)**
Supportive care	£481.40	1.79		£481.40	2.05	
Alitretinoin (30mg)	£3,337.49	2.01	£13,284.14	£3,337.49	2.16	£27,689.38
Analysis 4: Reinstate adverse events for alitretinoin only						
Treatment	**Cost**	**QALYs**	**ICER (per QALY)**	**Cost**	**QALYs**	**ICER (per QALY)**
Supportive care	£481.40	1.79		£481.40	2.05	
Alitretinoin (30mg)	£3,370.37	2.00	£14,072.21	£3,370.37	2.15	£29,199.56

QALY = quality-adjusted life-years; ICER = incremental cost-effectiveness ratio

Where patients are assumed to relapse into either the severe or moderate PGA state, it can be seen that the ICER for alitretinoin versus supportive care has risen from £13,432 (in the base case reanalysis) to £14,526 per QALY. This is because in the original model patients resumed treatment in the state they are instead expected to reach four weeks later (see chapter 5); correcting this error to reinstate the first four weeks of each subsequent treatment cycle and having patients relapse into the severe or moderate states reduces the utility associated with alitretinoin and increases the costs, raising its ICER versus supportive care.

Restricting the analysis to only those women who are potentially child-bearing raises the ICER (as expected, due to the costs of pregnancy testing, contraception and more regular dermatological visits) although this increase is not particularly large. Similarly, restricting the analysis to men only does not lower the ICER considerably.

Where adverse events are reinstated for alitretinoin only, the ICER for alitretinoin versus supportive care rises (as expected), although the increase is relatively modest – from £13,432 to £14,072 per QALY.

As noted in section 5.2.3, the ERG calculated that including prescription and monitoring costs for patients in remission at weeks 4 and 8 of each treatment cycle would increase the cost of alitretinoin by approximately £528 per patient over the course of three years. This cost has not been included in any of the additional analyses, but the impact of including this cost on each of the ICERs can be calculated by dividing £528 by the incremental utility in each case; for example, including this cost in the base case reanalysis with the BAP0003 utility would raise the ICER by approximately £528 / (2.01-1.79) = £2,400, from £13,432 to £15,832 per QALY.

Augustin utility data

The direction of the ICER movement in each analysis is identical with the Augustin utility data as with the BAP0003 data, although the magnitude of each ICER movement is greater. Furthermore, the ICERs for alitretinoin versus placebo are in the range £27,689 to £29,864 per QALY – slightly more than double the cost per QALY than under the alternative utility data.

7 Discussion

7.1 Summary of clinical effectiveness issues

The manufacturer's submission incorporated a full systematic review of the literature of the effects of alitretinoin in severe CHE refractory to topical steroid treatment. It is likely that all of the relevant evidence was identified in this review. The main findings are derived from trial BAP00089, which appears to be a generally well-conducted placebo controlled RCT. The manufacturer's interpretation of these findings appeared to be largely appropriate.

Though the included trials were of generally good quality, some issues around study validity were identified. A relatively high proportion of patients recruited to BAP00089 withdrew from the study due to either adverse events or insufficient response. This meant that only a minority of patients from BAP00089 were enrolled in BAP00091, raising the possibility of patient self-selection. This may partially explain the high rate of response (including to placebo) in trial BAP00091.

Another issue relating to BAP00091 is the apparent inconsistency in its outcome measures. While response rates as measured by PGA and PaGA were similar in the phase II and III trials, in BAP00091 there was poor agreement between these two different measures of treatment response, introducing further uncertainty around the true effects of retreatment with alitretinoin following relapse.

There were some instances of certain data not being reported in the submission (for example, the proportion of placebo patients in remission) and what appeared to be transcription errors (for example, clearly incorrect range values - reported as 95% CIs - for median time to relapse). However, these omissions were generally minor and in most cases did not prevent the ERG from appraising the reported results.

Crucially, given the chronic recurring nature of CHE, there is insufficient evidence on the efficacy and safety of alitretinoin beyond 48 weeks. Longer term follow-up is required to detect potentially rare adverse events and possibly to characterise the cardiovascular risks posed by the observed increase in cholesterol levels associated with alitretinoin treatment.

None of the evidence presented in the manufacturer's submission directly observed the clinical effect of alitretinoin treatment on health-related quality of life (HRQL). Instead, the manufacturer examined the relationship between HRQL (as measured by the DLQI) and disease severity (as measured by PGA) in patients with CHE. Though both of the available studies reported an association between the HRQL and PGA severity, one study implied a much stronger relationship than the other. If the effect of alitretinoin on HRQL is only to be indirectly estimated through its observed effect on PGA, then evidence for the degree of relationship between these outcomes needs to be more robust and consistent.

The main observed effects of alitretinoin were relative to placebo treatment with additional emollients where required. Though placebo-controlled trials are required for licensing purposes, and are not unethical given the steroid-refractory patients population, in practice some "steroid refractory" patients are likely to continue topical steroid treatment, even where the benefits are minimal. It remains unknown to what extent alitretinoin is effective relative to emollients and topical corticosteroids combined (the current first-line treatment choice).

To be included for treatment in the main trial presented here, patients had to formally be diagnosed as "severe" on the Physician's Global Assessment (PGA) outcome measure. However, in clinical practice, the most important factor in determining appropriateness of treatment would typically be the impact of the condition on the patient's activities, quantified by PGA or quality of life score such as DLQI. Hence in clinical practice patients qualifying for treatment with alitretinoin may well equate to the 'moderate' state in terms of the PGA. Though there is some evidence from BAP00003 that a predominately 'PGA moderate' CHE population will respond to alitretinoin treatment to some extent, there is no evidence for the effects of the licensed 30mg dose in this population.

The main trial BAP00089 was not powered to consider sub-groups and, in particular, the 'pompholyx only' group is very small, so no definitive conclusions about the effects of alitretinoin on subgroups should be drawn from this trial. However, the potential trend observed in terms of reducing hyperkeratosis fits with the effects of retinoids in other skin diseases and may warrant further investigation.

A change in threshold for the definition of 'relapse' from 75% to 50% of baseline mTLSS substantially reduced the time to 'relapse' observed in the 30 mg alitretinoin group. The manufacturer argued that the less stringent relapse definition of 50% mTLSS would yield a population with moderately severe CHE for whom the immediate retreatment with systemic therapy would be unlikely. However, clinical advice to the ERG suggests that in chronic skin disease, re-treatment is a different scenario from the first treatment. If the patient has had benefit and not experienced adverse effects then both they and their dermatologist are likely to consider re-treatment at lower levels of disease severity. Though this does not directly impact on the efficacy of alitretinoin in terms of achieving remission, it has clear implications for cost-effectiveness in terms of treatment-free time before beginning retreatment.

There is an almost complete lack of relevant good quality evidence for comparators specified in the scope for the treatment of CHE. The limited available data on PUVA and ciclosporin cannot be reasonably used to derive an estimate of their efficacy relative to alitretinoin. Though azothioprine is used in clinical practice, neither the manufacturer nor the ERG identified any published evidence on the efficacy of this treatment in patients with severe CHE.

In summary, though the evidence presented indicates that alitretinoin is efficacious in for the treatment of severe CHE, it gives little indication of alitretinoin's efficacy relative to likely alternative treatment options, or its efficacy and safety in the longer term.

7.2 Summary of cost effectiveness issues

In the manufacturer's submission to NICE, the base case ICERs reported for alitretinoin were £8614 per QALY versus ciclosporin, -£469 per QALY versus PUVA (with alitretinoin dominant) and £10,612 per QALY versus azathioprine. Relative to the base case, over longer time horizons these ICERs fell (random variation notwithstanding). In patients with hyperkeratotic CHE and in women of child-bearing potential, these ICERs rose slightly, but remained under £20,000. Where the utility values used in the model were replaced with those derived from an alternative study, these ICERs rose significantly (to £22,312 per QALY for alitretinoin versus azathioprine).

The submission had major shortcomings. The efficacy data for treatments other than alitretinoin were based on expert clinical opinion only. While the use of expert opinion may be justified where trial data do not exist to inform the relevant parameters, it should be elicited in a methodologically rigorous manner. The ERG remains unconvinced that this elicitation process generated reliable estimates of the efficacy of each of the comparator treatments.

The manufacturer assumed that patients receiving alitretinoin visited the dermatologist every four weeks and ceased treatment as soon as they responded, even if this was after only four or eight weeks of treatment. If in practice patients would receive treatment for longer than this then the manufacturer's model will have significantly underestimated the costs to the NHS.

Serious issues remain around the implementation of the model in Excel. The manufacturer has made substantial use of sparsely annotated VBA code and does not appear to have implemented a number of the assumptions given in the written submission correctly. In particular, the first four weeks of every subsequent treatment cycle are omitted. The definition of relapse used in the model does not correspond to that used in the relevant clinical trials. As a consequence the estimated costs and health outcomes presented by the manufacturer may be regarded as unreliable. The ERG has attempted to amend the VBA code in places to provide more appropriate estimates of the ICERs but in some cases this was not feasible.

Furthermore, the model originally submitted to NICE did not include a "supportive care" (or "placebo") arm and the treatment effects for alitretinoin were not placebo adjusted; as such, the model did not address whether alitretinoin was a cost-effective alternative to supportive care. Consequently, the ERG does not regard the ICERs generated by the manufacturer's original model as providing a reliable indication of the cost-effectiveness of alitretinoin compared to each of the comparators considered.

In response to a request from the ERG, the manufacturer provided a revised model with a "placebo" arm, and the comparison of alitretinoin with placebo made in this revised model is of greater merit given the more reliable efficacy data in the comparator arm. In this analysis, alitretinoin was reported to have an ICER of £12,931 per QALY gained versus placebo. However, the omission of adverse events entirely from this revised model, in combination with a number of other factors,

means that the model underestimates the costs of treatment associated with alitretinoin and so the true ICER may be higher. Additional analyses undertaken by the ERG produced ICERs close to £30,000 per QALY gained for alitretinoin versus supportive care. The model was deterministic and so cannot provide an estimate of the decision uncertainty associated with the results. There remains considerable uncertainty as to the true ICER of alitretinoin versus the relevant treatment comparators.

7.2.1 Implications for research

Given the limited duration of the available evidence, longer-term follow-up of trials or the implementation of registries are required to better establish the longer term efficacy and safety of alitretinoin.

Evidence of the effect of treatment on HRQL in patients with CHE is extremely limited. Future studies of alitretinoin should include a relevant HRQL measure (such as the DLQI) alongside measures of therapeutic response.

The placebo-controlled trials conducted to date have established that alitretinoin can be efficacious for the treatment of severe CHE refractory to topical steroids. However, future studies may want to establish the efficacy of alitretinoin relative to current first-line treatment (emollients plus topical steroids) and other treatments which are used in this indication (PUVA, azathioprine, ciclosporin).

8 References

1. Diepgen TL, Agner T, Aberer W, Berth-Jones J, Cambazard F, Elsner P, et al. Management of chronic hand eczema. *CONTACT DERMATITIS* 2007;57(4):203-210.
2. Agner T, Andersen KE, Brandao FM, Bruynzeel DP, Bruze M, Frosch P, et al. Hand eczema severity and quality of life: a cross-sectional, multicentre study of hand eczema patients. *Contact Dermatitis* 2008;59(1):43-7.
3. Van Coevorden AM, Williams HC, Svensson Å, Diepgen TL, Elsner P, Coenraads PJ. Interventions for hand eczema. (Protocol). *Cochrane Database of Systematic Reviews 2002, Issue 3. Art. No.: CD004055. DOI: 10.1002/14651858.CD004055*.
4. Basilea Pharmaceuticals. *DATA ON FILE reference sheet: Augustin Quality of life in patients with hand eczema - unpublished abstract. ACADEMIC IN CONFIDENCE. Reference number DOF-ALI08018. created 11/12/08*.
5. Petering H, Breuer C, Herbst R, Kapp A, Werfel T. Comparison of localized high-dose UVA1 irradiation versus topical cream psoralen-UVA for treatment of chronic vesicular dyshidrotic eczema. *JOURNAL OF THE AMERICAN ACADEMY OF DERMATOLOGY* 2004;50(1):68-72.
6. Sezer E, Etikan I. Local narrowband UVB phototherapy vs. local PUVA in the treatment of chronic hand eczema. *Photodermatology, Photoimmunology & Photomedicine* 2007;23(1):10-4.
7. Rosen K, Mobacken H, Swanbeck G. Chronic eczematous dermatitis of the hands: a comparison of PUVA and UVB treatment. *ACTA DERMATO-VENEREOLOGICA* 1987;67(1):48-54.
8. Simons JR, Bohnen IJ, van der Valk PG. A left-right comparison of UVB phototherapy and topical photochemotherapy in bilateral chronic hand dermatitis after 6 weeks' treatment. *CLINICAL & EXPERIMENTAL DERMATOLOGY* 1997;22(1):7-10.
9. Sheehan-Dare RA, Goodfield MJ, Rowell NR. Topical psoralen photochemotherapy (PUVA) and superficial radiotherapy in the treatment of chronic hand eczema. *BRITISH JOURNAL OF DERMATOLOGY* 1989;121(1):65-9.
10. Van Coevorden AM, Coenraads PJ, Svensson A, Bavinck JNB, Diepgen TL, Naldi L, et al. Overview of studies of treatments for hand eczema-the EDEN hand eczema survey. *BRITISH JOURNAL OF DERMATOLOGY* 2004;151(2):446-51.

11. Adams S, Bayerl C. Medium-dose-UVA-1 irradiation - and topical PUVA - Therapy in chronic dyshidrotic hand dermatitis - A prospective randomized study. [German]. *Aktuelle Dermatologie* 2007;33(4):142-145.
12. Grattan CE, Carmichael AJ, Shuttleworth GJ, Foulds IS. Comparison of topical PUVA with UVA for chronic vesicular hand eczema. *ACTA DERMATO-VENEREOLOGICA* 1991;71(2):118-22.
13. Woolacott N, Hawkins N, Mason A, Kainth A, Khadjesari Z, Vergel YB, et al. Etanercept and efalizumab for the treatment of psoriasis: a systematic review. *Health Technology Assesment* 2006;10(46):1-233.
14. Diepgen TL, Andersen KE, Brandao FM, Bruze M, Bruynzeel DP, Frosch P, et al. Hand eczema classification: a cross-sectional, multicentre study of the aetiology and morphology of hand eczema. *British Journal of Dermatology* 2009;160(2):353-8.
15. Sheehan MP, Atherton DJ, Norris P, Hawk J. Oral psoralen photochemotherapy in severe childhood atopic eczema: an update. *British Journal of Dermatology* 1993;129(4):431-6.
16. Schiener R, Gottlober P, Muller B, Williams S, Pillekamp H, Peter RU, et al. PUVA-gel vs. PUVA-bath therapy for severe recalcitrant palmoplantar dermatoses. A randomized, single-blinded prospective study. *Photodermatology, Photoimmunology & Photomedicine* 2005;21(2):62-7.
17. Grundmann-Kollmann M, Behrens S, Peter RU, Kerscher M. Treatment of severe recalcitrant dermatoses of the palms and soles with PUVA-bath versus PUVA-cream therapy. *Photodermatology, Photoimmunology & Photomedicine* 1999;15(2):87-9.
18. Engin B, Oguz O. Evaluation of time-dependent response to psoralen plus UVA (PUVA) treatment with topical 8-methoxypsoralen (8-MOP) gel in palmoplantar dermatoses. *INTERNATIONAL JOURNAL OF DERMATOLOGY* 2005;44(4):337-9.
19. Shephard SE, Schregenberger N, Dummer R, Panizzon RG. Comparison of 8-MOP aqueous bath and 8-MOP ethanolic lotion (Meladinine) in local PUVA therapy. *DERMATOLOGY* 1998;197(1):25-30.
20. Moon Jung K, Yoo Won C, Hae Young C, Ki Bum M. Comparison of local bath-PUVA with steroid treatment in palmoplantar pustular psoriasis and dyshidrotic eczema. [Korean]. *Korean Journal of Dermatology* 2000;38(6):742-749.
21. van Coevorden AM, Kamphof WG, van Sonderen E, Bruynzeel DP, Coenraads P-J. Comparison of oral psoralen-UV-A with a portable tanning unit at home

vs hospital-administered bath psoralen-UV-A in patients with chronic hand eczema: an open-label randomized controlled trial of efficacy. *ARCHIVES OF DERMATOLOGY* 2004;140(12):1463-6.

Appendix 1: Detailed critique of literature searches

Clinical-effectiveness searches

The submission appendices gave detailed descriptions of the search strategies and were designed to meet NICE requirements. They included the specific databases searched; the service providers used; the dates when searches were conducted; the date spans of the searches; the complete strategies used; the number of records identified for each search set; and the final result number. The search strategies were devised using comprehensive subject indexing and free text search terms; subject indexing was exploded whenever possible; and search facets were combined using Boolean operators. An RCT search filter was used and search results were restricted to humans. The date spans ran from database inception to October 2008.

The ERG has access to the same host providers as those used by the manufacturer, and has been able to reproduce the searches. The ERG was unable to replicate the search in EMBASE: the date span used by the manufacturer is broader than that available to the ERG. The EMBASE searches in the manufacturer submission go back to 1974 for the clinical evidence searches (and 1947 for the cost effectiveness searches). The ERG only has access to EMBASE from 1980.

Reproduction of the searches raised a number of issues.

It was reported that the latest search was carried out on 22nd October 2008. This does not explain why the date span for EMBASE was 1974 to 2008 week 24 (summer 2008). This might simply be a case of the latest search not being recorded: rerunning of the search would indicate this scenario.

The searches of MEDLINE, MEDLINE In-Process and EMBASE were run concurrently using the Ovid interface. Searches of MEDLINE In-Process are not directly reproducible as it is not possible to search on a specific named date, in this case the submission reported that the search was undertaken on 21st October 2008. More of a problem was the use of subject indexing terms in the strategy: these are redundant in MEDLINE In-Process as In-Process records have yet to be indexed. The search set for PUVA therapy (line 35) used subject indexing terms (but not free text terms). Therefore PUVA studies will not have been identified in a search of MEDLINE In-Process.

The use of subject indexing when searching across databases can be problematic. For instance, the MEDLINE Medical Subject indexing term (MeSH) for ‘Azathioprine/’

was used and this corresponds with an equivalent EMBASE Emtree term. However, the Emtree term for 'Alitretinoin/' was not used. This was not too great a problem as the free text term for 'alitretinoin' used the 'mp' suffix which includes the subject indexing field when searching.

Some effort was made to include Emtree as well as MeSH terms, but not consistently. For example, the Emtree for 'Hand disease/' was used as an equivalent subject indexing term to the MeSH 'Hand dermatoses/', but the Emtree term 'Hand eczema/' was not used. This term would have been searched for as it appears in the 'Dermatitis/' tree and this term was exploded. The MeSH term 'Ultraviolet therapy/' was used, but not the equivalent Emtree term 'Ultraviolet radiation/'. In most cases the MeSH term used did have an equivalent term in Emtree, e.g. 'exp Dermatitis/', 'exp Eczema/', 'exp Phototherapy/'. To effectively search MEDLINE and EMBASE concurrently, the strategy should contain all relevant MeSH terms, plus the equivalent Emtree terms.

A number of subject indexing terms were 'exploded' when 'explosion' was not available. A number of lower subject indexing terms were included in the strategy despite exploded terms higher in their tree being included. This had no detrimental effect on the search strategy, but rendered the terms redundant. For example 'exp Immunosuppressive agents/' includes 'Azathioprine/', 'Cyclopsorins/' and 'Cyclosporine/' in its tree, and 'exp Phototherapy/' includes 'Photochemotherapy/', 'PUVA therapy/' and 'Ultraviolet therapy/' in its tree. Therefore the individual lower terms could have been omitted. Searching for free text terms using 'mp' should have compensated for any missing or misused subject index terms.

Some of the search lines were unusual. In the search line where the comparator terms were combined (line 38) two lines were inadvertently left out (lines 30 and 37). Fortunately this did not impact on the final search results. Searches in the publication type field (pt) (line 22) included a number of unusual terms: 'trial' was searched for, as was 'clinical trial', 'controlled clinical trial', 'randomized controlled' – these subsequent terms were redundant. 'Multicentre' was included when this English spelling does not appear in either the MEDLINE or EMBASE publication type field. Search line 13 for '(treat$ or therap$).mp' in combination with 'hand eczema' terms was potentially too restrictive. The Cochrane review protocol search strategy (on which this strategy was based) used this line in order to restrict the results of looking for all possible interventions for hand eczema. Introducing named interventions rendered this line redundant.

Of more concern, three of the included comparator studies would not have been identified in the Ovid MEDLINE and EMBASE searches: Sheehan[15], Petering[5] and Simons[8]. All three studies are indexed in both MEDLINE and EMBASE. The Simons[8] study could have been identified in the search of PubMed, but neither of the Sheehan[15] or Petering[5] studies would have been. Furthermore, the Sheehan[15] study would never have been identified in searches for 'hand eczema' as hand (or any of the hand related search terms used) does not appear in the title, abstract or indexing of the records in either MEDLINE or EMBASE. It is accepted practice that studies are identified via checking of reference lists or handsearching. However, these methods were not reported in the submission.

It was not reported which issue of the Cochrane Library was searched. The results given are 33 records retrieved. Rerunning the search in the Cochrane Library (2008:issue 4) identified 122 records. It would appear that the search of the Cochrane Library only took into account the results from the CDSR and not the other three databases to be searched as instructed by NICE (CENTRAL, DARE and HTA). As the main purpose of the searches was to identify trials, it would have been useful to have searched the Cochrane trials register, CENTRAL. Further, of the 33 reviews identified in CDSR none included the Cochrane review protocol specifically about treatment for hand eczema on which the search strategies used in the submission were based[3].

PubMed was searched, but it is not clear why as MEDLINE had already been searched in the Ovid interface. There is no harm in doing this, but it is duplication of effort for no particular reason. However, in this instance it was worthwhile running a separate search in PubMed as one of the three included PUVA therapy studies unidentified in the Ovid interface searches was identifiable with the PubMed search strategy (Simons[8]).

It was reported that no additional searches were carried out. It is unclear how a number of references used in the submission were identified: guidelines, quality of life (QoL) studies and QoL indexes, epidemiological studies, ongoing trials (unpublished data), and conference presentations, abstracts and posters.

Cost-effectiveness searches

The databases searched for the cost effectiveness literature included MEDLINE, MEDLINE In-Process and EMBASE as required by NICE, but also DARE and HTA

which are not required. HEED was not searched, but use of this database is being reviewed at present. This is a subscription only database and the manufacturers may not have had access.

The date span for EMBASE was given as 1947-2008 week 46. It was not clear whether or not this was mistyped as the date span for the clinical evidence was 1974-2008. The date span for the CRD databases (NHS EED, DARE and HTA) was not reported, but this information is not easily identifiable from the CRD database website.

There was inconsistent use of field tags throughout this strategy. The 'mp' suffix was used for most search lines, but not for the set of terms used in the economic facet where 'ti,ab'(title and abstract field) was used. This means that the subject indexing terms used (all MeSH in this case) did not work in EMBASE. Using the 'mp' suffix with all free text terms in the clinical evidence search strategy ensured that the subject indexing fields (both MeSH and Emtree) were being searched.

Line 27 used 'cost$' truncated and then included 'costs or costly or costing', all of which were redundant. Of more importance was the inclusion of the term 'effective$' in this line. This term should not have been included here as it had a considerable impact on the number of records retrieved and on the relevance of those additional records retrieved: they were not cost/economic studies.

Duplicate records were not removed from the results of this search as they were from the clinical evidence searches.

The CRD databases search was very rudimentary as it used one term only; 'eczema'. Better practice would have been to have translated the MEDLINE/EMBASE strategy and included more terms (intervention/comparator terms, dermatitis, etc.). However, the explanation of how the results were hand searched, and how studies were subsequently retrieved or excluded, was helpful.

ERG search strategy

Clinical evidence.

MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (OvidSP). 1950-2008/Oct week 2. 18th December 2008.

	Searches	Results
1	clinical trial.pt.	460981
2	randomized.ab.	177355
3	placebo.ab.	111337
4	randomly.ab.	128722
5	trial.ab.	184716
6	groups.ab.	890308
7	dt.fs.	1318399
8	or/1-7	2462048
9	Hand Dermatoses/	5691
10	exp Dermatitis/	71158
11	exp Skin Diseases, Eczematous/	46755
12	(eczema$ or excema$ or dermat$ or tyloti$ or pompholyx or cheiropompholyx).ti,ab.	96347
13	(contact or allergic or irritant).ti,ab.	198833
14	(pulpitis or pulpite or dyshidro$ or dyshydro$ or dishidro$ or dishydro$ or hyperkerato$ or kerato$).ti,ab.	33323
15	or/10-14	339130
16	exp Hand/	58952
17	(hand$ or palm$ or finger$ or wrist$ or acra$ or apron or dors$).ti,ab.	497692
18	or/16-17	520568
19	15 and 18	20127
20	9 or 19	24110
21	Tretinoin/	16344
22	(retinoid$ or retinoic$).ti,ab.	27437
23	(alitretinoin or panretin or panretyn or panrexin or toctino).ti,ab,rn.	503
24	(5300-03-8 or 5352-74-9).rn.	499
25	exp Immunosuppressive Agents/	204067
26	(immunosuppress$ or immuno suppress$).ti,ab.	76781
27	exp Phototherapy/	20480
28	(puva or ultraviolet A or ultra violet A or UVA or UVB or ultraviolet B or ultra violet B or NBUVB or BBUVB or PNBUVB or REPUVA).ti,ab.	10781
29	(phototherap$ or photo therap$ or photochemotherap$ or photo chemotherap$ or photo chemo therap$).ti,ab.	5742
30	exp Cyclosporins/	32106
31	(cyclosporin$ or ciclosporin$ or csa or neoral or csaneoral or cya or cyc-a or	48565

	sandimmun$).ti,ab,rn.	
32	Azathioprine/	11889
33	(azathioprine or azothioprine or imuran or immuran or imurel or aza).ti,ab,rn.	21040
34	or/21-33	320939
35	8 and 20 and 34	605
36	Humans/	10826325
37	35 and 36	575
38	("20081117" or "20081114" or "20081016" or "20081030" or "20081106" or "20081024" or "20081031" or "20081027" or "20081112" or "20081103" or "20081020" or "20081023" or "20081111" or "20081105" or "20081107" or "20081110" or "20081017" or "20081021" or "20081022" or "20081029" or "20081118" or "20081104" or "20081015" or "20081028" or "20081113" or "20081114" or "20081117" or "20081118").ed.	105035
39	37 not 38	570

Trials Filter:
Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.0 (updated February 2008). The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org

EMBASE (OvidSP). 1980-2008/week 24. 18th December 2008.

	Searches	Results
1	random.tw.	84865
2	clinical trial.mp.	546865
3	exp Health Care Quality/	766213
4	or/1-3	1236716
5	Hand Disease/	1954
6	exp Dermatitis/	56018
7	Occupational Eczema/	1381
8	(eczema$ or excema$ or dermat$ or tyloti$ or pompholyx or cheiropompholyx).ti,ab.	81956
9	(contact or allergic or irritant).ti,ab.	160510
10	(pulpitis or pulpite or dyshidro$ or dyshydro$ or dishidro$ or dishydro$ or hyperkerato$ or kerato$).ti,ab.	24971
11	or/6-10	270579
12	exp Hand/	22598
13	(hand$ or palm$ or finger$ or wrist$ or acra$ or apron or dors$).ti,ab.	404502
14	or/12-13	409820
15	11 and 14	18012
16	5 or 15	19664
17	Alitretinoin/	1324
18	(retinoid$ or retinoic$).ti,ab.	25918

19	(alitretinoin or panretin or panretyn or panrexin or toctino).ti,ab,rn.	49
20	(5300-03-8 or 5352-74-9).rn.	1324
21	exp Immunosuppressive Agent/	307699
22	(immunosuppress$ or immuno suppress$).ti,ab.	67135
23	exp Phototherapy/	24016
24	(puva or ultraviolet A or ultra violet A or UVA or UVB or ultraviolet B or ultra violet B or NBUVB or BBUVB or PNBUVB or REPUVA).ti,ab.	10370
25	(phototherap$ or photo therap$ or photochemotherap$ or photo chemotherap$ or photo chemo therap$).ti,ab.	5061
26	Cyclosporin/	41416
27	(cyclosporin$ or ciclosporin$ or csa or neoral or csaneoral or cya or cyc-a or sandimmun$).ti,ab.	41952
28	Azathioprine/	45523
29	(azathioprine or azothioprine or imuran or immuran or imurel or aza).ti,ab.	13911
30	or/17-29	400765
31	4 and 16 and 30	457
32	Animal/ or Animal Experiment/ or Nonhuman/	3389412
33	(rat or rats or mouse or mice or murine or rodent or rodents or hamster or hamsters or pig or pigs or porcine or rabbit or rabbits or animal or animals or dogs or dog or cats or cow or bovine or sheep or ovine or monkey or monkeys).ti,ab,sh.	2234348
34	32 or 33	3601133
35	exp Human/ or Human Experiment/	6396073
36	34 not (34 and 35)	2941897
37	31 not 36	454
38	("200845" or "200843" or "200842" or "200831" or "200836" or "200827" or "200832" or "200828" or "200840" or "200826" or "200833" or "200839" or "200835" or "200837" or "200849" or "200838" or "200850" or "200846" or "200825" or "200830" or "200848" or "200844" or "200847" or "200834" or "200829" or "200841").em.	303862
39	37 not 38	418

Trials Filter:
Wong SS, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE. Journal of the Medical Library Association 2006;94(1):41-7.

CDSR (Cochrane Library). 2008:issue 4. 18th December 2008.

ID	Search	Hits
#1	MeSH descriptor **Hand Dermatoses** explode all trees	172
#2	MeSH descriptor **Dermatitis** explode all trees	2164
#3	MeSH descriptor **Skin Diseases, Eczematous** explode all trees	1634
#4	(eczema* or excema* or dermat* or tyloti* or pompholyx or cheiropompholyx):ti,ab,kw	7170

#5	(pulpitis or pulpite or dyshidro* or dyshydro* or dishidro* or dishydro* or hyperkerato* or kerato*):ti,ab,kw	1774
#6	**(**#2 OR #3 OR #4 OR #5**)**	9048
#7	MeSH descriptor **Hand** explode all trees	1506
#8	(hand* or palm* or finger* or wrist* or acra* or apron or dors*):ti,ab,kw	17169
#9	(#7 OR #8)	17235
#10	(#6 AND #9)	522
#11	(#1 OR #10)	522

NB 13 reviews were identified in CDSR

CENTRAL, DARE and HTA (Cochrane Library). 2008:issue 4. 18th December 2008.

ID	Search	Hits
#1	MeSH descriptor **Hand Dermatoses** explode all trees	172
#2	MeSH descriptor **Dermatitis** explode all trees	2164
#3	MeSH descriptor **Skin Diseases, Eczematous** explode all trees	1634
#4	(eczema* or excema* or dermat* or tyloti* or pompholyx or cheiropompholyx):ti,ab,kw	7170
#5	(pulpitis or pulpite or dyshidro* or dyshydro* or dishidro* or dishydro* or hyperkerato* or kerato*):ti,ab,kw	1774
#6	(#2 OR #3 OR #4 OR #5)	9048
#7	MeSH descriptor **Hand** explode all trees	1506
#8	(hand* or palm* or finger* or wrist* or acra* or apron or dors*):ti,ab,kw	17169
#9	(#7 OR #8)	17235
#10	(#6 AND #9)	522
#11	(#1 OR #10)	522
#12	MeSH descriptor **Retinoids** explode all trees	1490
#13	(retinoid* or retinoic*):ti,ab,kw	803
#14	(alitretinoin or panretin or panretyn or panrexin or toctino):ti,ab,kw	10
#15	MeSH descriptor **Immunosuppressive Agents** explode all trees	11781
#16	(immunosuppress* or "immuno suppress*"):ti,ab,kw	5347
#17	MeSH descriptor **Phototherapy** explode all trees	1335
#18	(puva or "ultraviolet A" or "ultra violet A" or UVA or UVB or "ultraviolet B" or "ultra violet B" or NBUVB or BBUVB or PNBUVB or REPUVA):ti,ab,kw	893
#19	(phototherap* or "photo therap*" or photochemotherap* or "photo	1416

	chemotherap*" or "photo chemo therap*"):ti,ab,kw	
#20	MeSH descriptor **Cyclosporins** explode all trees	2285
#21	(cyclosporin* or ciclosporin* or csa or neoral or csaneoral or cya or cyc-a or sandimmun*):ti,ab,kw	4359
#22	MeSH descriptor **Azathioprine** explode all trees	942
#23	(azathioprine or azothioprine or imuran or immuran or imurel or aza):ti,ab,kw	1828
#24	(#12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23)	19227
#25	(#11 AND #24)	105

NB 102 records were identified in CENTRAL, 0 records in DARE and 0 records in HTA

Economic evidence.

MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (OvidSP). 1950-2008/Nov week 2. 18th December 2008.

	Searches	Results
1	economics/	25938
2	exp "costs and cost analysis"/	142628
3	"Value of life"/	5091
4	economics, dental/	1800
5	exp "economics, hospital"/	15962
6	economics, medical/	7383
7	economics, nursing/	3861
8	economics, pharmaceutical/	2012
9	(economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.	277121
10	(expenditure$ not energy).ti,ab.	12011
11	(value adj1 money).ti,ab.	11
12	budget$.ti,ab.	12042
13	or/1-12	390480
14	((energy or oxygen) adj cost).ti,ab.	2067

15	(metabolic adj cost).ti,ab.	503
16	((energy or oxygen) adj expenditure).ti,ab.	11335
17	or/14-16	13359
18	13 not 17	387403
19	Hand Dermatoses/	5691
20	exp Dermatitis/	71158
21	exp Skin Diseases, Eczematous/	46755
22	(eczema$ or excema$ or dermat$ or tyloti$ or pompholyx or cheiropompholyx).ti,ab.	96347
23	(contact or allergic or irritant).ti,ab.	198833
24	(pulpitis or pulpite or dyshidro$ or dyshydro$ or dishidro$ or dishydro$ or hyperkerato$ or kerato$).ti,ab.	33323
25	or/20-24	339130
26	exp Hand/	58952
27	(hand$ or palm$ or finger$ or wrist$ or acra$ or apron or dors$).ti,ab.	497692
28	26 or 27	520568
29	25 and 28	20127
30	19 or 29	24110
31	Tretinoin/	16344
32	(retinoid$ or retinoic$).ti,ab.	27437
33	(alitretinoin or panretin or panretyn or panrexin or toctino).ti,ab,rn.	503
34	(5300-03-8 or 5352-74-9).rn.	499
35	exp Immunosuppressive Agents/	204067
36	(immunosuppress$ or immuno suppress$).ti,ab.	76781
37	exp Phototherapy/	20480
38	(puva or ultraviolet A or ultra violet A or UVA or UVB or ultraviolet B or ultra violet B or NBUVB or BBUVB or PNBUVB or REPUVA).ti,ab.	10781
39	(phototherap$ or photo therap$ or photochemotherap$ or photo chemotherap$ or photo chemo therap$).ti,ab.	5742
40	exp Cyclosporins/	32106

41	(cyclosporin$ or ciclosporin$ or csa or neoral or csaneoral or cya or cyc-a or sandimmun$).ti,ab,rn.	48565
42	Azathioprine/	11889
43	(azathioprine or azothioprine or imuran or immuran or imurel or aza).ti,ab,rn.	21040
44	or/31-43	320939
45	18 and 30 and 44	15
46	("20081113" or "20081118" or "20081117" or "20081114").ed.	22136
47	45 not 46	15

EMBASE (OvidSP). 1980-2008/week 46. 18th December 2008.

	Searches	Results
1	Health Economics/	10380
2	exp Economic Evaluation/	101056
3	exp Health Care Cost/	103597
4	exp PHARMACOECONOMICS/	55037
5	or/1-4	195699
6	(econom$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.	227367
7	(expenditure$ not energy).ti,ab.	9603
8	(value adj2 money).ti,ab.	443
9	budget$.ti,ab.	8669
10	or/6-9	235935
11	5 or 10	336962
12	(metabolic adj cost).ti,ab.	380
13	((energy or oxygen) adj cost).ti,ab.	1689
14	((energy or oxygen) adj expenditure).ti,ab.	9808
15	or/12-14	11386
16	11 not 15	334433
17	Hand Disease/	1954
18	exp Dermatitis/	56018
19	Occupational Eczema/	1381
20	(eczema$ or excema$ or dermat$ or tyloti$ or pompholyx or cheiropompholyx).ti,ab.	81956
21	(contact or allergic or irritant).ti,ab.	160510
22	(pulpitis or pulpite or dyshidro$ or dyshydro$ or dishidro$ or dishydro$ or hyperkerato$ or kerato$).ti,ab.	24971
23	or/18-22	270579

24	exp Hand/	22598
25	(hand$ or palm$ or finger$ or wrist$ or acra$ or apron or dors$).ti,ab.	404502
26	or/24-25	409820
27	23 and 26	18012
28	17 or 27	19664
29	Alitretinoin/	1324
30	(retinoid$ or retinoic$).ti,ab.	25918
31	(alitretinoin or panretin or panretyn or panrexin or toctino).ti,ab,rn.	49
32	(5300-03-8 or 5352-74-9).rn.	1324
33	exp Immunosuppressive Agent/	307699
34	(immunosuppress$ or immuno suppress$).ti,ab.	67135
35	exp Phototherapy/	24016
36	(puva or ultraviolet A or ultra violet A or UVA or UVB or ultraviolet B or ultra violet B or NBUVB or BBUVB or PNBUVB or REPUVA).ti,ab.	10370
37	(phototherap$ or photo therap$ or photochemotherap$ or photo chemotherap$ or photo chemo therap$).ti,ab.	5061
38	Cyclosporin/	41416
39	(cyclosporin$ or ciclosporin$ or csa or neoral or csaneoral or cya or cyc-a or sandimmun$).ti,ab.	41952
40	Azathioprine/	45523
41	(azathioprine or azothioprine or imuran or immuran or imurel or aza).ti,ab.	13911
42	or/29-41	400765
43	16 and 28 and 42	56
44	("200850" or "200847" or "200849" or "200848").em.	47609
45	43 not 44	54

NHS EED (CRD databases). 1994-2008/11. 18th December 2008.

#	1	eczema* OR excema* OR dermat* OR tyloti* OR pompholyx OR cheiropompholyx	474
#	2	contact OR allergic OR irritant	9596
#	3	pulpitis OR pulpite OR dyshidro* OR dyshydro* OR dishidro* OR dishydro* OR hyperkerato* OR kerato*	53
#	4	#1 or #2 or #3	10004
#	5	hand* OR palm* OR finger* OR wrist* OR acra* OR apron OR dors*	3551
#	6	#4 and #5	511
#	7	retinoid* OR retinoic*	20
#	8	alitretinoin OR panretin OR panretyn OR panrexin OR toctino	1
#	9	immunosuppress* OR “immuno AND suppress*”	281
#	10	puva OR “ultraviolet AND A” OR “ultra AND violet AND A” OR UVA OR UVB OR “ultraviolet AND B” OR “ultra AND violet AND B” OR NBUVB OR BBUVB OR PNBUVB OR REPUVA	114

#	11	phototherap* OR “photo AND therap*” OR photochemotherap* OR “photo AND chemotherap*” OR “photo AND chemo AND therap*”	60
#	12	cyclosporin* OR ciclosporin* OR csa OR neoral OR csaneoral OR cya OR cyc-a OR sandimmun*	195
#	13	azathioprine OR azothioprine OR imuran OR immuran OR imurel OR aza	107
#	14	#7 or #8 or #9 or #10 or #11 or #12 or #13	569
#	15	#6 and #14	13

NB 3 records were identified in NHS EED, none of which were relevant (10 records were identified in DARE and 0 in HTA)

Appendix 2: Justification for inclusion/exclusion of PUVA studies in manufacturer's submission

Study	Design, Control Type	Comparison	Justification for inclusion/exclusion in the analysis	
4 controlled PUVA studies – included in analysis				
Petering et al. 2003[5]	Controlled (within patient) trial	UVA-1 or topical PUVA	Controlled study, correct comparator - Chronic hand eczema (subtypes)	
Sezer et al. 2007[6]	Open label randomised, within-patient trial	UVB vs topical PUVA	Controlled study, correct comparator - Chronic hand eczema (subtypes)	
Rosen et al. 1987[7]	Open label, randomised controlled trial	UVB and oral PUVA with untreated hand controls.	Controlled study, correct comparator - Chronic hand eczema (subtypes)	
Simons et al. 1997[8]	Open-label randomised within-patient study	UVB and topical bath PUVA	Controlled study, correct comparator - Chronic hand eczema	
4 controlled PUVA studies – considered for, but not included in, analysis				
Sheehan-Dare et al. 1989[9]	Double-blind randomised within-patient study	PUVA and superficial radiotherapy	Controlled study, correct comparator - Chronic hand eczema	Patient data not recorded, only mean reduction in severity/extent of disease
Van Coevorden et al. 2004[10]	Open-label, randomised, controlled study	Oral and bath PUVA	Controlled study, correct comparator - Chronic hand eczema	Patient data not recorded, only mean reduction in severity/extent of disease
Adams et al. 2007[11]	Randomised (within-patient) study	PUVA and UVA-1.	Controlled study, correct comparator - Chronic hand eczema	Patient data not recorded, only mean reduction in severity/extent of disease
Grattan et al. 1991[12]	Double-blind randomised within -	Topical PUVA with UVA	Controlled study, correct comparator - Chronic hand eczema	Patient data not recorded, only mean

	patient trial			reduction in severity/extent of disease
Excluded studies				
Schiener et al. 2005 [16]	Randomised, single-blind, prospective	Bath PUVA versus gel-PUVA		Results not separated for hand and foot
Grundmann-Kollmann et al. 1999[17]	Randomised, controlled (within-patient) study	Bath-PUVA versus cream-PUVA		Results not separated for hand and foot
Engin et al. 2005[18]	Controlled (within-patient) study	Topical PUVA versus UVA		Results not separated for hand and foot
Shephard et al. 1998[19]	Controlled (within-patient) study	Bath PUVA versus lotion PUVA		Inadequately controlled
Jim et al. 2000[20]	Controlled study	Bath PUVA versus topical and oral steroid		Study mainly concerned with palmoplantar pustular psoriasis

Appendix 3: Summary of four PUVA studies meeting inclusion criteria for meta-analysis

Study	Comparison	Study population and severity scoring system	Results and safety
Petering et al. 2003 N=27[5]	UVA vs. topical PUVA.	All patients had recurrent disabling bilateral symmetrical vesicular hand eczema for at least 3 months with periods of remission not exceeding 2 weeks. DASI score10 for PUVA group at baseline (maximum 60).	DASI scores decreased significantly and were reduced to nearly half of the pre-treatment values in both arms (in the PUVA arm from 10 to 5 after 3 weeks and a further slight reduction at 6 weeks $p<0.05$). No significant differences between UVA and PUVA were detected. After 3 weeks no relapse was observed in 23 of 27 patients. Both treatments were well tolerated.
Sezer et al. 2007 N=12[6]	Comparison of paint-PUVA on one hand and UVB on the other hand.	Subtype only CHE of dry and dyshidrotic types, (hyperkeratotic CHE excluded). Erythema, squamation, induration, fissures and itching assessed on a scale (none 0 to severe 3). The total clinical score was the sum of each variable (max 15). Mean total clinical scores: UVB – 10.5, PUVA – 9.83.	Significant ($p<0.05$) reductions in total clinical scores for both treatments over a 9-week assessment period. 8% (1 patient) cleared and 75% marked clinical improvement with PUVA (9 patients), no improvement in 17% (2 patients). For PUVA mean clinical scores ($p<0.05$ compared to baseline): Baseline 9.83 ± 2.95, week three 8.50 ± 2.39. week six 5.42 ± 2.19, week nine 2.42 ± 1.44. Both treatments were considered equally effective. AEs mild xerosis observed in both groups and hyperpigmentation in the PUVA group. At 10 weeks follow up, 8 of 12 patients relapse free with UVB and 6 of 12 relapse free with PUVA.
Rosen et al. 1987 N=35[7]	Oral PUVA (N=18) and UVB (N=17). One hand exposed the other an untreated control. .	Bilateral hand eczema, symmetrical distribution and severity of at least 6 months duration. Predominantly females (31/35) with vesicular CHE (26/31) enrolled. Two patients were hyperkeratotic in the PUVA arm. Clinical assessment of: desquamation, erythema, vesiculation, infiltration and fissures. Each variable was assessed on a four point scale: 0, none; 1, slight; 2, moderate; 3, severe (range 5-18).	For PUVA mean severity score before treatment 10.6 ± 0.8 in the treated hand and 10.6 ± 0.8 in the untreated hand. After treatment 0.8 ± 0.2 in the treated hand and 5.4 ± 0.7 in the untreated hand. PUVA: 92% reduction in severity score at treatment cessation. 14 patients cleared (4 patients at 3 weeks, 5 patients at 6 weeks and 5 patients at 9 weeks, $p<0.001$) In 9/14 PUVA patients dermatitis recurred within 3 months (mean) of end of treatment. PUVA was considered to be superior to UVB. AEs in PUVA were nausea, oedema, pain and itching in the treated hand, hyperpigmentation, soreness and stiffness in the fingers.
Simons et al. 1997 N=13[8]	UVB vs. topical PUVA on each hand.	Patients with vesicles or hyperkeratotic plaques of the hands present for > 6 months. Clinical assessment score (based upon area and severity of symptoms) from baseline to 6 weeks. Mean severity score for PUVA was 10.17 ± 2.26.	Mean severity scores reduced to 7.66 for PUVA after 6 weeks, a 25% reduction for PUVA treatment. There was no significant improvement between the treatment modalities. Six patients suffered phototoxic reactions from PUVA on a total of 9 occasions. The PUVA treated side became more pigmented than the UVB treated side. Relapse not assessed.
Sheehan-Dare et al. 1989 N=25[9]	UVA vs. PUVA	All patients had chronic eczematous changes on the palms for at least 6 months with either continuous or episodic vesiculation. Clinical severity score graded 0 (normal skin) to 4 (active pompholyx).	Significant improvements in clinical severity scores from baseline to 6, 9 and 18 weeks in both groups (p value not reported). Mean scores reduced to between 2-3 at 6, 9 and 18 weeks. Significantly better clinical improvement seen for UVA

		Mean clinical severity score at baseline 3-4.	Relapse not reported.
Van Coevorden et al. 2004 N=158[21]	Oral PUVA at home N=78, Hospital administered bath PUVA N=80.	Chronic bilateral or unilateral hand eczema of at least 1 year's duration, at least 2 relapses or more than 3 consecutive weeks with visible signs in the last 3 months and moderate to severe hand eczema with a hand eczema severity score of at least 6 on a 0-21 based on the sum of severity ratings (0-3). Mean severity score at baseline 8.1 in both groups.	At week 10: Oral PUVA mean score 4.8 (95% CI, 3.9-5.6) (mean reduction 3.3), bath PUVA mean score 5.6 (95% CI, 4.7-6.4) (mean reduction 2.5). In the oral PUVA group 72% improved and in the bath PUVA group 61% improved. At 8 weeks follow up scores did not change significantly. 23% and 18% in oral and bath groups respectively worsened by more than 1 point. Efficacy of both treatments was comparable. Relapse not reported.
Adams et al. 2007 N=15[11] (*Paper in German)*	One hand received topical PUVA and the other medium-dose UVA-1	Chronic dyshidrotic hand eczema.	At 5 weeks: Significant improvement in DASI (dishydrotic eczema area and severity index) score with PUVA (p=0.0498; n=11). There was no significant difference between the two therapies (p=0.3070), both of which resulted in a significant decrease in DASI. Relapse not reported in abstract.
Grattan et al. 1991 N=15[12]	One hand received topical PUVA and the other UVA.	Recurrent disabling bilateral symmetrical vesicular hand eczema for at least 6 months with periods of remission not exceeding 1 month in the previous 6. At baseline: Mean severity score <2.5 (clear 0 - severe 4) for both treatments; mean VAS score (0-10) 0 between 3 and 5 for both treatments; mean T_{120} score (area and severity scoring system, 0-120) 27.63 (PUVA) and 26.63 (UVA).	The reduction in severity score was significant (p<0.005) for both PUVA and UVA-treated hands after 8 weeks treatment on the T_{120} score and global rating scales but only for the UVA-treated hand on the VAS. A further reduction in severity scores during the 8-week follow-up period did not reach statistical significance on any of the assessment tools. No significant change in any scores at 4 or 8 weeks after end of treatment. Relapse not reported.

Appendix 4: Summary of ciclosporin trial

Study ID	Number of Subjects by Treatment Arm Entered	Baseline disease demographics/ severity scoring system	Efficacy Results
Granlund et al. 1996 N=41	Ciclosporine compared with BDP cream; patients assigned to either over 6 weeks.	Patients with hand eczema continuously for at least 6 months, causing significant disability and who had an inadequate response to conventional treatment.	Disease activity score decreased to 57% of baseline (12.9 to 7.3) in the ciclosporine group (mean change -6, SD 4.3; $p<0.001$). Relapses occurred to the same extent in both groups. After a 2-week follow-up 50% of patients in both groups had relapsed. Adverse events occurred in 68% of the ciclosporine group.

Appendix 5: Validation of manufacturer's VBA code

The Excel model's main VBA module contains four treatment Sub procedures – *treatment1, treatment2, treatment3* and *treatment4* – each corresponding to a different treatment comparator. The code in each treatment Sub procedure is similar – each procedure contains a "main loop" which loops around once every treatment cycle, and this main loop is contained within a larger loop which loops around 100 times for each patient. The code forcing patients to enter or re-enter the severe state ("*currentState = 1*") is given at the start of the larger loop but is not given within the main loop – as such, when patients resume treatment following relapse they are not forced to re-enter the severe state.

Appendix 6: Amendments to the VBA code by the ERG for analysis 3

In each of the *treatment* Subs in the VBA code (discussed in chapter 5), the following code was inserted at the beginning of the "main loop":

```
If currentState = 4 And i = trtStartMonth - 1 Then
  If Rnd > 0.306 Then currentState = 1 Else currentState = 2
End If
```

In layman's terms, this reads: "If the patient is currently in remission (currentState = 4) and the start month of the next treatment cycle has been reached (i = trtStartMonth) then, if a randomly generated number between 0 and 1 is greater than 0.306 (i.e. with probability 69.4%) force the patient into the severe state (If Rnd > 0.306 Then currentState = 1), otherwise force the patient into the moderate state (Else currentState = 2)".

Chapter 5 also discussed the issue that the assumed 24 week time to relapse for alitretinoin appeared to be measured from 12 weeks into the treatment cycle for all patients, irrespective of the individual patient's time of entering remission – this was also the case for the supportive care arm. This assumption appears to be hard-coded into the model in the following VBA code, given in each of the relevant *treatment* Subs:

```
trtStartMonth = (i - 1) - 12 / 4 + (timeToRelapse + remissionTrtTime)
```

The ERG amended this code to allow patients to incur the utility associated with the severe or moderate PGA state for the first four weeks. This was achieved by modifying this code as follows:

trtStartMonth = i - 12 / 4 + (timeToRelapse + remissionTrtTime)