Ruxolitinib for treating non-segmental vitiligo in people 12 years and older [ID3998]:

A Single Technology Appraisal

Produced by

Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU

Authors

Will Sullivan[1,2]
Hollie Wheat[1,2]
Ash Bullement[1,2]
Sophie Robinson[1]
Alex Allen[1]
Viktoria Eleftheriadou[3]
G.J. Melendez-Torres[1]
Caroline Farmer[1]

[1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter

[2] Delta Hat Limited, Nottingham

[3] Walsall Healthcare NHS Trust & The Royal Wolverhampton NHS Trust

Correspondence to

Caroline Farmer
3.09 South Cloisters, St Luke's Campus, Heavitree Road, Exeter, EX1 2LU; c.farmer@exeter.ac.uk

Date completed 16/10/2023

Source of funding This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR136073.

Declared competing interests of the authors	None
Acknowledgments	The authors acknowledge the administrative support provided by Mrs Sue Whiffin and Ms Jenny Lowe (both PenTAG).
Rider on responsibility for document	The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors
This report should be referenced as follows	Sullivan, Wheat, Bullement, Robinson, Allen, Eleftheriadou, Melendez-Torres, Farmer. Ruxolitinib for treating non-segmental vitiligo in people 12 years and older [ID3998]: A Single Technology Appraisal. Peninsula Technology Assessment Group (PenTAG), 2023.

Author Contributions:	
Will Sullivan	Critical appraisal of the economic evidence and analysis submitted by the company, conducted additional economic analyses and drafted sections of the report
Hollie Wheat	Critical appraisal of the economic evidence and analysis submitted by the company, conducted additional economic analyses and drafted sections of the report
Ash Bullement	Critical appraisal of the economic evidence and analysis submitted by the company, conducted additional economic analyses and drafted sections of the report
Sophie Robinson	Critical appraisal of the company's literature search strategies and conducted additional literature searches as part the EAG appraisal
Alex Allen	Critical appraisal of the company's clinical effectiveness evidence and drafted sections of the report
Viktoria Eleftheriadou	Expert clinical advice to the EAG about NS vitiligo and its treatment
G.J. Melendez-Torres	Critical appraisal of the company submission, writing and editorial input
Caroline Farmer	Project lead, critical appraisal of the company submission, writing and editorial input

Table of Contents

Abbreviations 9

1. Executive summary 11
    - 1.1. Overview of the EAG's key issues 11
    - 1.2. Overview of key model outcomes 13
    - 1.3. The decision problem: summary of the EAG's key issues 14
    - 1.4. The clinical effectiveness evidence: summary of the EAG's key issues 15
    - 1.5. The cost effectiveness evidence: summary of the EAG's key issues 17
    - 1.6. Other key issues: summary of the EAG's views 21
    - 1.7. Summary of EAG's preferred assumptions and resulting ICER 21
2. Introduction and Background 23
    - 2.1. Introduction 23
    - 2.2. Critique of the company's description of the underlying health problem 23
    - 2.3. Critique of the company's overview of current service provision 24
    - 2.4. Critique of company's definition of decision problem 25
3. Clinical Effectiveness 35
    - 3.1. Critique of the methods of review(s) 35
    - 3.2. Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these) 36
        - 3.2.1. Studies included in the clinical effectiveness review 36
        - 3.2.2. Description and critique of the design of the studies 42
        - 3.2.3. Description and critique of the results of the studies 58
    - 3.3. Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison 73
    - 3.4. Additional work on clinical effectiveness undertaken by the EAG 73
        - 3.4.1. Network meta-analysis 74
        - 3.4.2. Unanchored matching-adjusted indirect comparison 75
    - 3.5. Conclusions of the clinical effectiveness section 76
4. Cost-effectiveness 78
    - 4.1. EAG comment on company's review of cost-effectiveness evidence 78
    - 4.2. Summary and critique of company's submitted economic evaluation by the EAG 78
        - 4.2.1. NICE reference case checklist 78
        - 4.2.2. Validation 79
        - 4.2.3. Population 80
        - 4.2.4. Interventions and comparators 81

4.2.5. Model structure and logic 85
4.2.6. Perspective, time horizon and discounting 88
4.2.7. Treatment effectiveness and extrapolation 90
4.2.8. Health-related quality of life 94
4.2.9. Resources and costs 99
5. Cost-effectiveness results 105
5.1. Company's cost-effectiveness results 105
5.1.1. Base case results 105
5.2. Company's sensitivity analyses 106
5.2.1. One-way sensitivity analysis 106
5.2.2. Probabilistic sensitivity analysis 106
5.2.3. Scenario analyses 107
5.3. Model validation and face validity check 110
6. External Assessment Group's Additional Analyses 111
6.1. EAG corrections and adjustments to the company's base case model 112
6.2. Exploratory and sensitivity analyses undertaken by the EAG 113
6.2.1. Alignment of management costs with clinical practice 113
6.2.2. Face validity of utility values 114
6.2.3. Ruxolitinib dosing 115
6.2.4. Duration of costs application in 'non-response' 116
6.2.5. Approach to handling missing data for no regain of response 116
6.2.6. Retreatment with ruxolitinib 116
6.2.7. Impact on the ICER of additional clinical and economic analyses undertaken by the EAG 117
6.3. EAG's preferred assumptions 119
6.3.1. Scenario analyses around the EAG's preferred assumptions 122
6.4. Conclusions of the cost-effectiveness section 125
References 126
Appendix A: EAG indirect treatment comparison feasibility analysis 129
Appendix B: Multi-way sensitivity analysis of utility values 132

List of key issues

Key Issue 1: The clinical and cost effectiveness of ruxolitinib as compared to established treatment options is unknown 14

Key Issue 2: The clinical effectiveness evidence presented by the company was not representative of the target population and the population used in the company's economic evaluation 16

Key Issue 3: Cost-effectiveness model's structural assumptions and use of clinical effectiveness data 17

Key Issue 4: Approach to ruxolitinib dosing assumptions in the cost-effectiveness model 17

Key Issue 5: Approach to resource use and cost assumptions in the cost-effectiveness model 18

Key Issue 6: Approach to patient utility assumptions in the cost-effectiveness model 19

Key Issue 7: Approach to adverse event assumptions in the cost-effectiveness model 20

List of tables

Table 1: Summary of key issues 11

Table 2: Key differences between the company's preferred assumptions and EAG's preferred assumptions 12

Table 3: Summary of EAG-preferred assumptions and tentative preferred cost-effectiveness results 21

Table 4: Company rationale and EAG view on the positioning of ruxolitinib and consideration of vehicle cream as the principal comparator 27

Table 5: Summary of decision problem 29

Table 6: Summary of EAG's critique of the methods implemented by the company to identify evidence relevant to the decision problem 35

Table 7: Completed clinical trials of ruxolitinib for the treatment for vitiligo 40

Table 8. Summary of exposure in TRuE-V1/TRuE-V2 (adapted from Table 37, CS Appendix C [SmPC report]) 50

Table 9: Clinical outcomes for ruxolitinib appraised by the EAG 52

Table 10. Quality assessment of the outcomes linked to the TRuE-V1/TRuE-V2 prior therapy subgroup 57

Table 11: Change in facial vitiligo outcomes based on pooled data from TRuE-V1 and TRuE-V2 61

Table 12: Physician- and Patient-reported improvement in facial vitiligo from TRuE-V1 and TRuE-V2 63

Table 13: Shift summary of maintenance response on F-VASI (ITT pooled population) from week 24 to week 52 (ruxolitinib arm) 64

Table 14: F-VASI75 at week 104 for those who responded in the TRuE-V1 and TRuE-V2 trials 65

Table 15: Change in bodily vitiligo outcomes based on pooled data from TRuE-V1 and TRuE-V2 67

Table 16: Shift summary of maintenance response on T-VASI (ITT pooled population) from week 24 to week 52 (ruxolitinib arm) 69

Table 17: NICE reference case checklist 78

Table 18: TRuE-V1[18] and TRuE-V2[19] ruxolitinib exposure summary statistics, adapted from Tables 6 and 7 of the company's response to EAG Clarification Q B10. 83

Table 19. Key TRuE-V1[18], TRuE-V2[19] and TRuE-LTE[9] data and assumptions applied in the company's analysis (CS Table 33) 90

Table 20: Reproduction of CS Table 37 – combined results informing the company's base case utility assumptions 96

Table 21: Absolute expected utility values implied by Table 18 and applied in the company's model 97

Table 22: Health state utility values applied in the company's deterministic analysis 97

Table 23: Company base case results 105

Table 24: Overview of company's scenario analyses (CS Table 46) 108

Table 25: Summary of key cost-effectiveness results from scenario analyses (CS Table 47) 108

Table 26: EAG-corrected company base case results 112

Table 27: EAG's exploratory analyses with EAG fixes applied 117

Table 28: From EAG-corrected company base case results to EAG-preferred tentative base case 1 results (all deterministic) 120

Table 29: Summary EAG-preferred tentative base case 1 results 120

Table 30: From EAG-corrected company base case results to EAG-preferred tentative base case 2 results (all deterministic) 121

Table 31: Summary EAG-preferred tentative base case 2 results 122

Table 32: Individual impact of each scenario upon the EAG's deterministic Base Case 1 ICER 123

Table 33: Individual impact of each scenario upon the EAG's deterministic Base Case 2 ICER 124

Table 34. EAG indirect treatment comparison feasibility analysis 129

Table 35: Multi-way sensitivity analysis of utility values (aligned to tentative EAG-preferred base case 1) – ICER for ruxolitinib versus vehicle cream 133

Table 36: Multi-way sensitivity analysis of utility values (aligned to tentative EAG-preferred base case 2) – ICER for ruxolitinib versus vehicle cream 133

List of Figures

Figure 1: Phase II trial study design 38

Figure 2: Participant flow in the pooled TRuE-V1 and TRuE-V2 trials, including the OLE and LTE 45

Figure 3: Change (mean ±SE) in F-VASI in TRuE-V1 and TRuE-V2 62

Figure 4: Change (mean ±SE) in T-VASI in TRuE-V1 and TRuE-V2 68

Figure 5: Company's model structure schematic (CS Figure 20) 85

Figure 6: Relationship between model time horizon and company's list price base case deterministic ICER 89

Figure 7: Kaplan-Meier treatment discontinuation projections stratified by TRuE study and treatment arm, as presented by the company in response to EAG clarification question B8. 93

Figure 8: PSA Scatterplot (CS Figure 22) 106

Figure 9: PSA scatterplot, EAG-preferred tentative base case 1 121

Figure 10: PSA scatterplot, EAG-preferred tentative base case 2 122

Abbreviations

A&E	Accident and emergency
AE	Adverse event
BAD	British Association of Dermatologists
BID	Twice daily
BNF	British National Formulary
BSA	Body surface area
BSC	Best supportive care
CI	Confidence interval
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical Study Report
DLQI	Dermatology life quality index
DP	Depigmentation
DSU	Decision Support Unit
EAG	External Assessment Group
EMA	European Medicines Agency
eMIT	Drugs and pharmaceutical electronic market information tool
EQ-5D	EuroQol five dimension
FA	Feasibility assessment
FS	Focal Seizures
GNX	Ganaxolone
F-BSA	Facial body surface area
F- PaGIC	Facial Patient Global Impression of Change
F-PhGVA	Facial Physician's Global Vitiligo Assessment
F-VASI	Facial Vitiligo Area Scoring Index
GP	General practitioner
HADS	Hospital Anxiety and Depression Scale
HR	Hazard ratio
HRQoL	Health-related quality of life
HSUV	Health State utility value
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
ITC	Indirect treatment comparison
ITT	Intention-to-treat
LTE	Long-term extension
MHRA	Medicines and Healthcare products Regulatory Agency
NA	Not applicable
NB-UVB	Narrowband ultraviolet B
NHS	National Health Service
NICE	National Institute for Health and Care Excellence

NMA	Network meta-analysis
NR	Not reported
OLE	Open label extension
NSV	Nonsegmental vitiligo
ONS	Office of National Statistics
OR	Odds Ratio
OWSA	One-way sensitivity analysis
PAS	Patient access scheme
PBO	Placebo
PSA	Probabilistic sensitivity analysis
QA	Quality assessment
QALY	Quality-adjusted life year
QD	Once daily
RCT	Randomised controlled trial
RPS	Repigmentation scores
SD	Standard deviation
SE	Standard error
SLR	Systematic literature review
SmPC	Summary of product characteristics
TA	Technology Appraisal
TEAE	Treatment emergent adverse events
TCI	Topical calcineurin inhibitors
TCS	Topical corticosteroids
T-VASI	Total Vitiligo Area Scoring Index
VASI	Vitiligo Area Scoring Index
VitiQoL	Vitiligo-specific quality-of-life instrument
VNS	Vitiligo noticeability scale
VS	Versus
WTP	Willingness to pay

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the external assessment group (EAG) as being potentially important for decision making. It also includes the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.6 explain the key issues in more detail. Background information on the condition, technology and evidence and information on non-key issues are in the main EAG report.

All issues identified represent the EAG's view, not the opinion of NICE.

1.1. Overview of the EAG's key issues

A brief overview of the key issues identified by the EAG in their appraisal of the company submission (CS) is provided in Table 1. Further detail of the issues is provided in Sections 1.3, 1.4, 1.5, and 1.6.

The EAG identified a key issue relating to the decision problem of the appraisal, in that the CS did not demonstrate the clinical and cost effectiveness of ruxolitinib in comparison to the relevant comparators in the company's proposed positioning. A further key clinical issue related to the absence of a comprehensive evidence base in the CS for the target population for ruxolitinib, defined by the company as those who have not responded to topical corticosteroids and/or calcineurin inhibitors, or for whom these treatments are contraindicated. In terms of cost effectiveness issues, the EAG noted key issues with the company's model structure and use of clinical effectiveness data, and key patient utility and healthcare cost assumptions. Owing to outstanding key issues, the EAG was only able to tentatively state preferred cost-effectiveness results.

Table 1: Summary of key issues

ID	Summary of issues	Report sections
Key Issue 1	The clinical and cost effectiveness of ruxolitinib as compared to established treatment options is unknown	2.4, 4.2.4
Key Issue 2	The clinical effectiveness evidence presented by the company was not	2.4, 3.2.2.2, 4.2.3

ID	Summary of issues	Report sections
	representative of the target population and the population used in the company's economic evaluation	
Key Issue 3	Cost-effectiveness model's structural assumptions and use of clinical effectiveness data	4.2.5, 4.2.7
Key Issue 4	Approach to ruxolitinib dosing assumptions in the cost-effectiveness model	4.2.4
Key Issue 5	Approach to resource use and cost assumptions in the cost-effectiveness model	4.2.9
Key Issue 6	Approach to patient utility assumptions in the cost-effectiveness model	4.2.8
Key Issue 7	Approach to adverse event assumptions in the cost-effectiveness model	4.2.8, 4.2.9

The key differences between the company's preferred assumptions and the assumptions in the tentative EAG-preferred analyses are outlined in Table 2.

Table 2: Key differences between the company's preferred assumptions and EAG's preferred assumptions

	Company's preferred assumption	EAG preferred assumption	Report Sections
Treatment pathway resource use	Assumed that most patients in the "non-response" state (after ruxolitinib or vehicle cream discontinuation) incurred ongoing active treatment and disease management in secondary care.	Considered a comparison to vehicle cream to only be potentially relevant for an end-of-line positioning and assumed far lower ongoing active treatment and disease management costs in the "non-response" state.	1.3, 1.5, 2.4, 4.2.3, 4.2.9, 6.2.1
Patient utility	Company's multistep approach produced utility values for "maintenance" and "stable" states that were higher than age-equivalent general population estimates. The company's model categorised those achieving F-VASI50-74 repigmentation	Preferred to limit health state utility values to be no greater than age-equivalent general population estimates and corrected the inconsistency of assigning non-response utility values to patients achieving F-VASI50-74 repigmentation.	1.5, 4.2.8, 6.2.2

	Company's preferred assumption	EAG preferred assumption	Report Sections
	improvements at 24 weeks as non-responders, despite the multistep approach predicting higher utility values.		
Expected ruxolitinib dose	Assumed that the median pooled daily dose of trial drug (ruxolitinib or vehicle cream) from the pooled TRuE-V dataset represented the expected daily dose of ruxolitinib in practice.	Used the mean ruxolitinib dose estimate from TRuE-V summary data to inform dose expectations. As this mean estimate was greater than the maximum recommended dose in the product licence for ruxolitinib, the EAG presented two alternative dosing approaches: one in which the cost of mean dose was assumed; another in which the cost of the maximum recommended dose was assumed.	1.5, 4.2.4, 6.2.4

Abbreviations: EAG, External assessment Group F-VASI, facial vitiligo area scoring index; F-VASI50-74, 50% to 74% improvement from baseline in F-VASI

1.2. Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is an estimate of the extra cost of every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Improving depigmentation caused by NS vitiligo, and thereby improving health-related quality of life

Overall, the technology is modelled to affect costs by:

- Adding acquisition costs of ruxolitinib to the treatment pathway
- Offsetting downstream costs, by predicting a treatment effect in delaying and reducing time spend in "non-response", incurring treatment and disease management costs

The modelling assumptions that have the greatest effect on the ICER are:

- The number of ruxolitinib tubes required for an average treatment course
- The patient utility values assumed to be associated with each modelled health state.
- The cost of downstream treatments and secondary care, especially in relation to the positioning of ruxolitinib as reflected by the company's model.

1.3. The decision problem: summary of the EAG's key issues

The EAG identified one key issue with regard to the decision problem for this appraisal.

Key Issue 1: The clinical and cost effectiveness of ruxolitinib as compared to established treatment options is unknown

Report sections	**2.4, 4.2.4**
Description of issue and why the EAG has identified it as important	The NICE decision problem for this appraisal was to evaluate ruxolitinib in comparison with established clinical management, which the EAG understood to be other topical treatments (including TCS and TCIs), NB-UVB therapy, betamethasone in those with rapidly progressing disease, and combinations of these as indicated. The company submission, including the company's economic evaluation, was based on a comparison between ruxolitinib and vehicle cream (i.e. a placebo therapy). The EAG considered that a comparison with vehicle cream was only relevant for the end of the treatment pathway; i.e. after all other treatment options have been considered. However, the company stated that the appropriate positioning for ruxolitinib would be at the 2nd line position, between the use of TCS/TCIs and NB-UVB therapy. Clinical advice to the EAG was also that a 2nd line position would be more appropriate for ruxolitinib. However, the EAG did not consider that the CS was consistent with this positioning.
What alternative approach has the EAG suggested?	At clarification (question B1), the EAG requested that the company re-formulate their economic evaluation to represent a specific position in the treatment pathway, i.e., to compare ruxolitinib with the existing treatment options that it would displace. The company declined to do this. The EAG has been unable to resolve this issue during its appraisal.
What is the expected effect on the cost-effectiveness estimates?	In principle, if the efficacy of the control arm of the model was increased to reflect the use of active treatment options, the magnitude of QALY gain may decrease compared to the company's base-case analysis, which would cause the ICER to increase. However, since treatment options would incur additional cost, the incremental costs would be expected to decrease, which would cause the ICER to decrease. It was not possible for the EAG to comment on the likely magnitude of effect on the ICER due to the infeasibility of robust comparisons using available evidence for alternative comparators and in consideration of the broader structural issues with the company's model (Key Issue 3).
What additional evidence or analyses might help to resolve this key issue?	Fundamentally, the EAG considered that the company should have conducted a head-to-head trial to compare ruxolitinib with the alternative treatment options at its proposed positioning and did not

Report sections	2.4, 4.2.4
	accept the company's rationale for this not being necessary or appropriate. With the existing evidence base, the EAG accepted arguments from the company that estimating the effectiveness of established treatment options relative to ruxolitinib was challenging, given heterogeneity in trial design used to evaluate treatment options for ruxolitinib. The company should have performed a narrative synthesis of evidence for the different treatment options, to consider the relative effectiveness of treatment options in consideration of variation in trial design, and the EAG did not accept the company's rationale for not doing this. However, this would only provide an insight into the potential effectiveness of ruxolitinib as compared to existing treatment options and would not have provided reliable effect estimates for use in economic modelling. At this point, the EAG considered that either (a) ruxolitinib be considered as a final treatment option only, after all other treatment options had been considered [thus the company's analysis is relevant] or (b) the company's analysis should be re-submitted using a reasonable estimate of effectiveness for the relevant treatment comparator.

Abbreviations: EAG, External Assessment Group; TCIs, topical calcineurin inhibitors; TCS, topical corticosteroids

1.4. The clinical effectiveness evidence: summary of the EAG's key issues

The EAG considered that the CS lacked a comprehensive overview of the clinical effectiveness evidence for ruxolitinib. Notably, several clinical trials of ruxolitinib that appeared relevant to the decision problem were not included in the CS, and clinical effectiveness evidence from the included trials was not fully presented within the CS Document B. The company provided data for the scoped outcomes in supplementary documents, such as the Summary of Product Characteristics Report (SmPC) for ruxolitinib produced by the European Medicines Agency (EMA) and in PDF documents from their clinical trial reports. However, these data were difficult to identify from the documents and not always presented in a form that could support a transparent appraisal by the EAG during the timeframe of the EAG. Overall, the EAG considered that the CS presented by the company undermined the ability of the EAG to conduct a full appraisal of the clinical effectiveness evidence for ruxolitinib. However, during its appraisal the EAG did not identify any indication that the lack of transparency in the CS would meaningfully effect cost effectiveness estimates. The EAG therefore did not make this issue one of its key issues.

The EAG identified one key issue with the clinical effectiveness evidence for ruxolitinib, which was related to the discrepancy between the evidence submitted by the EAG and their proposed target population for ruxolitinib.

Key Issue 2: The clinical effectiveness evidence presented by the company was not representative of the target population and the population used in the company's economic evaluation

Report sections	**2.4, 3.2.2.2, 4.2.3**
Description of issue and why the EAG has identified it as important	The company suggested that ruxolitinib should be positioned as a 2nd line treatment option, to be considered after TCS and TCIs. However, the clinical effectiveness evidence presented in the CS was based on the full trial populations, only 28% of whom had previously received TCS or TCI treatment. The EAG was uncertain whether clinical outcomes would be expected to differ according to the line of treatment received. In response to clarification (question A2), the company provided a series of documents containing tables with clinical data for the previously treated subgroup. The files were inadequately labelled and the format of the data prevented a thorough appraisal by the EAG within the timeline available, however the EAG noted a slightly higher response rate to ruxolitinib in those who had previously received treatment compared to the full trial population, as assessed using the F-VASI75. As the EAG had not received a full submission for this population (including population characteristics including the prevalence of effect modifiers) and could not compare this finding across outcomes, the EAG was unsure if this was evidence of a true difference in treatment effect between treatment lines. Clinical data from the previously treated subgroup (any previous treatment) were used in the company's economic model, but without a comprehensive and transparent submission of evidence for the previously treated subgroup, the EAG cannot validate if the use of these clinical data was appropriate.
What alternative approach has the EAG suggested?	At clarification (QA2), the company were invited to provide evidence for the prior treated subgroup, however this was submitted in a format that could not be appraised during the timeframe of the EAG appraisal. The EAG was unable to resolve this issue during its appraisal.
What is the expected effect on the cost-effectiveness estimates?	Since this was a fundamental issue concerning the scope of the appraisal, it was not possible for the EAG to comment on the potential effect on cost-effectiveness estimates.
What additional evidence or analyses might help to resolve this key issue?	To inform committee decision-making, the company should submit clinical effectiveness evidence for the previously treated subgroup in a transparent manner using the Document B template as a guide. This should include population and intervention characteristics for the subgroup (i.e., to demonstrate that clinical effects used in the model were reliable) and clinical outcome data for the subgroup across all scoped outcomes (to demonstrate if outcomes vary between the previously treated subgroup and the full population, and to validate the choice of clinical inputs used in the economic model).

Abbreviations: EAG, External Assessment Group F-VASI, facial vitiligo area scoring index; F-VASI75, 75% improvement from baseline in F-VASI; TCI, topical calcineurin inhibitors; TCS, topical corticosteroids

1.5. The cost effectiveness evidence: summary of the EAG's key issues

The EAG identified five key issues with the cost effectiveness evidence for ruxolitinib submitted by the company.

Key Issue 3: Cost-effectiveness model's structural assumptions and use of clinical effectiveness data

Report sections	4.2.5, 4.2.7
Description of issue and why the EAG has identified it as important	The company's chosen model structure assumed that patients who achieve F-VASI50-74 at ~24 weeks discontinue treatment owing to non-response. This is neither in line with expectations for clinical practice nor in line with the company's own registrational trials. Separately, the company has made questionable structural economic assumptions in the model. For example, it is structurally impossible in the company's model for a patient in the "Maintenance period" state with F-VASI75-89 to achieve F-VASI≥90 and therefore transition to the model's "Stable" state. Overall, the EAG considered the company's method to incorporate data from the TRuE-V trials into the model to be subject to substantial limitations. Ultimately, this meant that the EAG had little confidence in the results of the model.
What alternative approach has the EAG suggested?	The EAG-corrected company base case (Section 6.1) corrects for a calculation error in the company's model. Other structural corrections and re-specifications were not feasible during the timeframe of the EAG's appraisal but the EAG considered these to be important for robust decision-making.
What is the expected effect on the cost-effectiveness estimates?	The combined effect of correcting for structural errors and exploring structural assumptions and the use of data that better fit expected clinical practice is unclear. While these issues are outstanding, the EAG was unable to present more than tentative EAG-preferred results.
What additional evidence or analyses might help to resolve this key issue?	The company can address this issue by addressing the structural problems identified by the EAG and documented in Section 4.2.5, and otherwise respecify the model structure to reflect expected clinical practice in line with the EAG critique (Sections 4.2.5 and 4.2.7)

Abbreviations: EAG, External Assessment Group

Key Issue 4: Approach to ruxolitinib dosing assumptions in the cost-effectiveness model

Report sections	4.2.4
Description of issue and why the EAG has identified it as important	The company's analysis assumed that the median daily dose of trial drug (ruxolitinib or vehicle cream) from the pooled TRuE-V dataset was equivalent to the expected daily dose of ruxolitinib in practice. However, it would have been more appropriate to use the mean dose of ruxolitinib, rather than the median dose across arms. Moreover, as the TRuE-V dosing data were skewed to the right, the mean dose of ruxolitinib in the TRuE-V dataset was greater than the median and greater than the dose limit of two 100mg tubes per month specified in the product licence for ruxolitinib[1]. This was important as the expected

Report sections	4.2.4
	per-patient use (cost) of ruxolitinib was uncertain and a key driver of cost-effectiveness results.
What alternative approach has the EAG suggested?	The EAG used a mean ruxolitinib dose estimate from TRuE-V summary data provided by the company in response to clarification question B10 to inform dose expectations in its preferred analyses. As this mean estimate was greater than the maximum recommended dose in the product licence for ruxolitinib, the EAG presented two alternative dosing approaches: one in which the cost of mean dose was assumed, another in which the cost of maximum recommended dose was assumed. The difference between these approaches was the difference between EAG-preferred tentative base cases 1 and 2 (Section 6.3).
What is the expected effect on the cost-effectiveness estimates?	Using the mean TRuE-V ruxolitinib dose (or maximum recommended dose) as a proxy for the expected ruxolitinib dose increased the expected cost of ruxolitinib and increased the EAG-corrected company base case ICER by £82,412 (£58,260), as shown in Section 6.2.7.
What additional evidence or analyses might help to resolve this key issue?	Further TRuE-V dosing data beyond the summary data provided by the company in response to clarification question B10 would further clarify doses received across participants in the TRuE-V trials. Clinical and patient expert opinion on expected ruxolitinib use in in the NHS and the likelihood of the licenced maximum dose being exceeded would help build understanding of expected doses used in practice.

Abbreviations: EAG, External Assessment Group

Key Issue 5: Approach to resource use and cost assumptions in the cost-effectiveness model

Report sections	4.2.9
Description of issue and why the EAG has identified it as important	The EAG was concerned that the company overestimated disease management and subsequent treatment resource use in its economic analysis. With respect to Key Issue 1, the EAG considered that a comparison to vehicle cream was only potentially appropriate as an end-of-line comparison. In this instance, assuming that any dermatology outpatient attendances or NB-UVB treatment after ruxolitinib or standard of care treatment (as the company do in the "non-response state") would be inappropriate. The EAG was also concerned that the company's psychological support assumptions overestimated the proportion of patients who would receive NHS psychological support. Even if the company's positioning of ruxolitinib as a 2nd line treatment option could be considered appropriate, the EAG considered that the company's NB-UVB and dermatology attendance and psychological support assumptions overestimated resource use, in a manner that biased cost-effectiveness results in favour of ruxolitinib.
What alternative approach has the EAG suggested?	The EAG removed dermatology outpatient and NB-UVB costs from "non-response" health state costs and reduced the proportion of patients expected to receive psychological support in the EAG-preferred tentative base cases. Separately, the EAG explored

Report sections	4.2.9
	scenarios assuming different levels of dermatology outpatient engagement in the "non-response" health state.
What is the expected effect on the cost-effectiveness estimates?	Compared with the EAG-corrected company's base-case results, making these adjustments caused total costs across arms to decrease and the incremental cost associated with ruxolitinib to decrease. This change in isolation causes the EAG-corrected company base case ICER to increase by £85,603, as shown in Section 6.2.7.
What additional evidence or analyses might help to resolve this key issue?	Resolution of Key Issue 1 would be the first step in clarifying appropriate resource use assumptions for patients who are in a "non-response" state after ruxolitinib and the care it would displace. Following this, further clinical expert validation of resource use frequency assumptions would help further resolve uncertainty and potential bias in the company's assumptions.

Abbreviations: EAG, External Assessment Group; ICER, incremental cost effectiveness ratio; NB-UVB, narrow-band ultraviolet B therapy; NHS, National Health Service

Key Issue 6: Approach to patient utility assumptions in the cost-effectiveness model

Report sections	4.2.8
Description of issue and why the EAG has identified it as important	The company's approach to estimate utility values for the health states in their economic model was complex and subject to numerous important limitations and assumptions. Notably, the values generated lacked face validity, implying better-than-general-population utility for patients in "Maintenance" or Stable" states. Elsewhere, the company's assignment of utility values to health states was internally inconsistent given their own estimation procedure. The company estimated a utility value of 0.890 for patients achieving F-VASI50-74 at 24 weeks. Yet, in the company's model, patients achieving F-VASI50-74 were categorised as "non-responders" and assigned a utility value of 0.797.
What alternative approach has the EAG suggested?	In EAG-preferred tentative analyses, the EAG limited health state utility assumptions to be no greater than age-adjusted general population expectations and adjusted the utility value assumed for the "non-response" state to account for the proportion of TRuE-V ruxolitinib patients expected to have achieved F-VASI50-74 at 24 weeks (assumed in the company's analysis to be "non-responders"). Separately, the EAG conducted further health state utility scenario analyses to explore the sensitivity of results to different assumptions, given the uncertainty in the estimates produced by the company.
What is the expected effect on the cost-effectiveness estimates?	Compared with the EAG-corrected company base case, applying EAG-preferred adjustments reduced the incremental QALY gain predicted for ruxolitinib. These changes in isolation caused the EAG-corrected company base case ICER to increase by £12,188 to £8,006, as shown in Section 6.2.7
What additional evidence or analyses might help to resolve this key issue?	The EAG considered that the company should have assessed HRQoL in the trials of ruxolitinib using a validated generic HRQoL measure, such as the EQ-5D, particularly given limitations in the psychometric validation of the VitiQoL measure used in the TRuE-V trials. It was not clear that the company used their systematic review to identify the best

Report sections	4.2.8
	available data to inform utility assumptions but given the TRuE-V HRQoL data collected and issues with indirect comparisons cited in the CS, there may not be substantial additional published data to further resolve uncertainty. Further patient and clinical expert testimony could help further understanding of appropriate health state utility assumptions.

Abbreviations: EAG, External Assessment Group F-VASI, facial vitiligo area scoring index; F-VASI50-74, 50% to 74% improvement from baseline in F-VASI

Key Issue 7: Approach to adverse event assumptions in the cost-effectiveness model

Report sections	4.2.8, 4.2.9
Description of issue and why the EAG has identified it as important	The company's economic analysis did not account for the HRQoL implications of adverse events, despite treatment-emergent adverse events affecting 47.7% of ruxolitinib participants in the pooled TRuE-V population, as documented in Table 23 of the CS. Treatment-arm specific expectations for adverse event costs were captured using incidence rates of adverse events occurring in ≥4% of trial participants. Though ruxolitinib was a topical treatment with no clear safety concerns in registrational trials, the EAG was concerned that the company was introducing bias in favour of ruxolitinib. The EAG's concern was heightened if ruxolitinib was considered as an end-of-line treatment (Key Issue 1), and as such would replace no treatment (no toxicity) and given the evidence from TRuE-V data that some people may expose themselves to more ruxolitinib than indicated in the product licence (Key Issue 4). Further, the EAG noted a tendency in TRuE-V trials for some patients to use more than the recommended maximum dose of ruxolitinib, which could have safety implications if such a tendency was seen in practice despite label warnings. Lastly, the company-preferred analysis predicted a modest lifetime QALY gain associated with ruxolitinib ([redacted] QALYs), while tentative EAG-preferred estimates were more modest still (~[redacted] QALYs). When incremental QALY gain estimates are of this magnitude, it was plausible that accounting for the HRQoL implications of adverse events appropriately could meaningfully affect cost-effectiveness results.
What alternative approach has the EAG suggested?	In clarification question B16, the EAG asked the company to incorporate utility and cost implications for the adverse event data in Table 24 of the CS (treatment-emergent adverse events occurring in ≥ 1% of patients in any treatment group), and in doing so to report utility, resource and cost data identification methods, and justify any assumptions required in absence of data. In response, the company did not comply with the EAG's request, or alter their CS approach to account for adverse events in the cost-effectiveness analysis in any way.
What is the expected effect on the cost-effectiveness estimates?	If vehicle cream or no active treatment was considered an appropriate comparator (Key Issue 1), appropriately accounting for the expected cost and HRQoL effects of adverse events associated with ruxolitinib would increase the expected costs and reduce the expected QALYs associated with ruxolitinib, reducing its estimated cost-effectiveness.

Report sections	4.2.8, 4.2.9
What additional evidence or analyses might help to resolve this key issue?	The company can address this key issue by complying with the EAG's request in clarification question B16; specifically, to incorporate utility and cost implications for the adverse event data in Table 24 of the CS (treatment-emergent adverse events occurring in ≥ 1% of patients in any treatment group), and in doing so to report utility, resource and cost data identification methods, and justify any assumptions required in absence of data.

Abbreviations: CS, company submission; EAG, External Assessment Group; HRQoL, health-related quality of life; QALY, quality-adjusted life year

1.6. Other key issues: summary of the EAG's views

The EAG did not identify any further key issues.

1.7. Summary of EAG's preferred assumptions and resulting ICER

Table 3 summarises the corrections and EAG-preferred changes to the company base case analysis, and their isolated and collective implications for cost-effectiveness results. As described through Sections 1.1 to 1.5, several EAG concerns remain unresolved. As such, the EAG-preferred results shown are tentative only. EAG adjustments collectively reduce the expected incremental QALY gain associated with ruxolitinib while increasing its expected incremental cost, leading to EAG-preferred tentative ICERs that were far in excess of the relevant NICE decision-making threshold range.

Table 3: Summary of EAG-preferred assumptions and tentative preferred cost-effectiveness results

Scenario	Incremental cost	Incremental QALYs	ICER (change from company base case)
Company's base case	[redacted]	[redacted]	£13,634 (+£0)
All EAG fixes to correct the company's base case applied	[redacted]	[redacted]	£13,031 (-£603)
Disable NB-UVB & vehicle cream costs, set proportion of patients receiving psychological support to 15% and proportion of patients using dermatology resources in the no response health state to 0% to represent end of treatment pathway	[redacted]	[redacted]	£100,036 (+£86,402)
Utility values capped at general population in response health states, and 'no response' utility value set to weighted average of 'no response' and F-VASI50-74	[redacted]	[redacted]	£22,639 (+£9,005)

The pooled mean over both TRuE-V studies through week 1 to week 52 exclusively for the ruxolitinib arm*	[redacted]	[redacted]	£97,359 (+£83,725)
Maximum daily dose as specified in the product licence (stating that no more than two x 100g tubes per month should be used)**	[redacted]	[redacted]	£73,000 (+£59,366)
Patients in the 'no response' health state on the ruxolitinib arm still accrue drug acquisition and disease management for a lifetime horizon	[redacted]	[redacted]	£4,114 (-£9,520)
Assume missing data are for non-responders	[redacted]	[redacted]	£16,283 (+£2,649)
EAG tentative preferred Base Case 1	[redacted]	[redacted]	£303,189 (+£289,555)
EAG tentative preferred Base Case 2	[redacted]	[redacted]	£262,880 (+£249,246)

Abbreviations: EAG, External Assessment Group; eMIT, Drugs and pharmaceutical electronic market information tool; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year
*Applicable only to EAG Base Case 1
**Applicable only to EAG Base Case 2

Modelling errors identified and corrected by the EAG are described in Section 6.1. For further details of the exploratory and sensitivity analyses done by the EAG, see Section 6.2.

2. INTRODUCTION AND BACKGROUND

2.1. Introduction

In this report, the External Assessment Group (EAG) provides a review of the evidence submitted by Incyte for the appraisal of ruxolitinib for the treatment of non-segmental vitiligo (NSV) in people aged 12 years and older.

2.2. Critique of the company's description of the underlying health problem

NSV is a depigmented skin disorder characterised by acquired, progressive, and depigmented lesions of the skin, mucosa, and hair[2]. It is believed to be caused mainly by the autoimmune loss of melanocytes from the involved areas. It is frequently associated with other autoimmune diseases, particularly autoimmune thyroid diseases including Hashimoto's thyroiditis and Graves' disease, rheumatoid arthritis, type 1 diabetes, psoriasis, pernicious anaemia, systemic lupus erythematosus, Addison's disease, and alopecia areata.

NSV is an umbrella term, which encompasses most forms of vitiligo experienced by people. The company detail the classification of NSV in Table 3 (Doc B). The two classifications notable for this submission are 'generalised or common' where patches are often symmetrical, and can affect any part of the skin, mainly hands, fingers, face, and trauma-exposed areas, and 'acro-facial' where patches affect the face, head, hands, and feet, and typically involves the perioral region and the extremities of digits.

NSV tends to spread slowly with new patches developing off and on throughout a person's life. These patches may range from specks of depigmentation through to complete depigmentation. Development of vitiligo happens through what are termed "flare-ups" and flare-ups arise from an autoimmune attack on functional melanocytes. There is no known way to predict when a person will experience a flare-up, the location of the flare-up or which patch might flare, and how far it will spread. The EAG's clinical expert stated that this unpredictable spread is distressing for the person and can lead to considerable additional depigmentation.

The EAG's clinical expert noted that 20% to 30% of people with NSV have rapidly progressing disease. This could typically a spread from 5% of a person's body surface area (BSA) to 30% of their BSA over a period of two weeks. It is difficult to capture the proportion of people

experiencing this due to NHS waiting times. People who are identified as having rapidly progressing disease use an amended treatment pathway.

Section B.1.3.1 of the company submission (CS) provided an overview of NSV. Based on advice from the EAG's clinical expert, the CS presented an accurate overview of diagnosis and classification, clinical presentation, development, epidemiology and disease burden. The company also expanded on the humanistic burden of NSV for people with the condition and their carers.

The company noted that there was currently no consensus on the methods to assess the extent of a person's vitiligo but expand upon body surface area (BSA) and Vitiligo Area Scoring Index (VASI), the methods used in the pivotal trials of ruxolitinib. The VASI method combines the extent of depigmentation with the degree of pigmentation. The EAG's clinical expert confirmed that VASI scales are accurate, however they are not widely used in practice as they are time consuming to complete.

2.3. Critique of the company's overview of current service provision

There is no NICE guideline for the management of vitiligo. In Figure 4 of the CS, the company provided an overview of the recommendations from the British Association of Dermatologists (BAD) for the management of vitiligo. The EAG considered the recommendations as reported by the company to be broadly correct, and clinical advice to the EAG was that this pathway would be followed for both young people (aged 12 and over) and adults in the NHS.

The BAD guidelines specified that the specific treatment sequence received by people with vitiligo was influenced by the progressive nature of the condition, the extent of areas affected (body surface area; BSA), the specific areas of skin affected (e.g. sensitive skin around the eyes and genitals), the level of distress experienced by the person, individual preference, and a risk-benefit profile that considers the likely risks of treatment alongside the likely benefits for the individual. This means that while the pathway in the BAD guidelines was applicable to the target population, there was likely to be variation in what treatments people would receive. However, in the main, the EAG understood that topical corticosteroids (TCS) or calcineurin inhibitors (TCIs; e.g. tacrolimus) were typically received as first-line treatment, and that people may receive both of these in sequence. If TCS and/or TCIs are not effective or not indicated, people may be considered for narrowband ultraviolet B therapy (NB-UVB) alone or in combination with either TCS or TCIs. A large UK-based trial[3] demonstrated that NB-UVB therapy in combination with

topical steroids was more effective for vitiligo than either treatment alone, however clinical advice was that NB-UVB therapy was still used more commonly than the combination treatment, though this may change with time and further dissemination of the BAD guidelines. Topical treatments may be less preferable for those whose condition has a high BSA, as would hand-held NB-UVB therapy as opposed to 'full cabinet' NB-UV therapy. Clinical advice to the EAG was that handheld NB-UVB was only available in a small number of centres in the UK, and so most people using this treatment would travel to healthcare settings to receive full cabinet NB-UVB. There were no specific management recommendations for those with vitiligo affecting the face, except that some existing treatments may not be considered appropriate for the sensitive area around the eyes. In those with rapidly progressive disease (approximately 30% of the population), people may not be administered TCS or TCIs and instead may initially be offered NB-UVB therapy. Oral betamethasone would be the second line option for those with rapidly progressive disease. Other treatments for vitiligo mentioned in the BAD guideline, such as de-pigmentation treatments and surgery, were not routinely available in the NHS.

Clinical expert advice to the EAG was that waiting times for referrals to a dermatologist may be long, and so people may receive a short course of TCS from their GP while waiting for specialist input. At that time point they may receive a longer course of TCS and/or TCIs or would be considered for other treatment options. There may be a high level of attrition within the treatment pathway, as people with vitiligo stop pursuing (further) active treatments and use maintenance strategies only, such as camouflage make-up and sunscreen. This may be particularly true for treatments such as full-cabinet NB-UVB therapy that require multiple in-person appointments, as these can be challenging to schedule alongside school and work.

Clinical advice to the EAG was that people with NSV who are experiencing psychological distress related to their condition would be encouraged to self-refer to psychological services. Self-help techniques may also be recommended for those with mild distress.

2.4. Critique of company's definition of decision problem

The population for this appraisal specified in the NICE scope was 'people aged 12 years and older with NSV with facial involvement', which was consistent with the marketing authorisation for ruxolitinib. The company proposed that treatment should be limited to people who have not responded to topical treatments (TCS or TCI) or for whom topical treatments are contraindicated, not tolerated or otherwise medically inadvisable. In principle, the EAG did not disagree with the company's proposed positioning after 1st line, and clinical advice to the EAG

was that topical treatments would still be considered before ruxolitinib as a treatment option. However, the EAG considered it plausible that the availability of another treatment in the 2nd line position may increase the number of people who choose to move beyond TCS and TCI, either to pursue further efficacy gains and/or because of concerns about the potential side effects of TCS and TCI treatment.

Consistent with their proposed positioning of ruxolitinib, between current 1st and 2nd line treatment options, the company argued that the relevant comparator for ruxolitinib was the comparator used in their clinical trials (vehicle cream, i.e. a placebo) as there was no other treatment currently in this position of the pathway. At clarification (question B1), the EAG requested that the company provide evidence to substantiate their positioning of ruxolitinib, particularly with reference to the company's argument that the comparator in their clinical trials (vehicle cream) would be the relevant comparator in its economic analyses. An overview of the response provided by the company and the EAG appraisal of this is shown in Table 4. Overall, the EAG considered that the relevant comparator for this position in the treatment pathway was existing 2nd line treatment options, which for most people with vitiligo will be NB-UVB therapy with or without TCS or TCIs. The EAG understood that those with rapidly progressing disease would not be ineligible for treatment with ruxolitinib, and so betamethasone may also be a relevant comparator at 2nd line. While clinical advice to the EAG was consistent with evidence presented by the company that many people with vitiligo were not receiving any active treatment for their condition, the EAG did not consider that this negated the need for this appraisal to determine the clinical and cost effectiveness of ruxolitinib relative to existing treatments used in the NHS. Moreover, the EAG did not consider that the company had provided evidence or rationale to conclude that the same factors influencing treatment use would not also affect ruxolitinib (and so ruxolitinib would not be a realistic treatment option for those not currently receiving treatment). However, the EAG considered that a comparison with no treatment may be a reasonable comparator to ruxolitinib in the 3rd line position in the treatment pathway, as at this point in the pathway there were no existing treatment options available (and therefore a comparison with no treatment would reflect the choice for people with vitiligo at this stage of treatment). Clinical advice to the EAG was consistent with the company's proposed positioning of ruxolitinib as a 2nd line treatment option. The EAG therefore did not disagree with the company's proposed positioning, but rather considered that the evidence base submitted by the company was not appropriate for decision-making in this position. This issue is outlined in Key Issue 1.

Table 4: Company rationale and EAG view on the positioning of ruxolitinib and consideration of vehicle cream as the principal comparator

Company response to clarification	EAG view
A retrospective cohort study amongst vitiligo patients in the UK found that among the prevalent cohort of 44,910 patients in 2019, 85.0% of patients were not on vitiligo-related treatment. In the first year after diagnosis, 60.8% of patients did not receive any vitiligo-related treatment (e.g., topical steroids, topical calcineurin inhibitors, oral steroids, phototherapy), increasing to ≥82.0% from the second year onward[4]. This finding is indicative of the vast majority of prevalent patients, including those with prior failure with TCS or TCI, not proceeding to another line of off-label therapy. In the first year, patients were recorded as having been prescribed topical corticosteroids (29.1%), topical calcineurin inhibitors (11.8%), and oral corticosteroids (4.2%). From the second year onward, the percentage of patients prescribed oral corticosteroids remained stable, while prescription of topical corticosteroids and calcineurin inhibitors declined to 11.4% and 3.9% in the second year, respectively, remaining low thereafter[4].	Consistent with the evidence presented by the company, clinical advice to the EAG was that many people with vitiligo may not be receiving treatment. Clinical advice to the EAG was that this may be due to frustration with long waiting lists to see a consultant about their condition. As ruxolitinib was expected to be prescribed by a consultant, the EAG considered it plausible that uptake of ruxolitinib would be similarly affected, and that ruxolitinib would not therefore offer an alternative treatment option for this group. Clinical advice was also that people with vitiligo may not receive treatment due to a lack of effective treatment options. The EAG assumed that this may affect treatment uptake after existing treatment options had been exhausted (i.e. after participants had not responded to 2nd line treatment options). In this context, 2nd line treatment options would still be relevant for comparison with ruxolitinib. The EAG considered it plausible that there may be people not receiving treatment for their condition because 2nd line treatment options were contraindicated, or collaborative decision-making between clinician and patient had determined that the balance of risks and benefits were not acceptable. The EAG therefore considered that the availability of ruxolitinib would encourage some people in this group to seek treatment when they would not otherwise, though the precise numbers of people this would affect are unknown. Overall, the EAG did not consider that the company had presented sufficient evidence or rationale to determine whether a group not receiving treatment would do so following the availability of ruxolitinib.
Given the availability of generic TCS and TCI, ruxolitinib cream is not anticipated to be cost-effective in the full population.	The EAG agreed that, given the widespread availability of TCS and TCIs used in the 1st line and the evidence that a significant minority of people with vitiligo respond well to these, it was likely that ruxolitinib would not be cost effective for use as a 1st line treatment for this population.
This positioning is considered most appropriate since introduction of a topical treatment after failure of initial topical treatment but prior to phototherapy is less burdensome for patients with vitiligo and less of a strain on NHS resources.	The EAG considered that the resource use associated with ruxolitinib as compared with NB-UVB therapy could be more appropriately considered within a cost effectiveness analysis comparing these treatments.
There remains a lack of equitable access to phototherapy, which is further compounded by other competing chronic inflammatory skin disease	As noted above, the EAG was aware that many people with vitiligo may not receive active therapies due to difficulties in accessing care. However, the

indications for phototherapy such as psoriasis and atopic dermatitis, resulting in long wait times and variability in receiving this treatment option across the UK	EAG did not consider that this negates the need to determine the clinical and cost effectiveness of ruxolitinib relative to available alternative treatments. Moreover, the EAG did not consider that the company had provided evidence or rationale to determine whether ruxolitinib would be used by people who were not accessing existing treatments. For example, and as noted above, the EAG understood that ruxolitinib would be prescribed by a consultant, and therefore may not be received by people who do not seek treatment from a consultant.
Clinicians generally recommend that phototherapy is prioritised for patients with large BSA (i.e., >10%) affected[5,6].	This issue raised by the company was consistent with clinical advice to the EAG that those with a larger BSA of vitiligo may find topical treatments less pragmatic, and so may prefer to receive NB-UVB (phototherapy). However, the EAG was aware that NB-UVB may be administered through the use of a hand-held device, suitable for smaller areas of the body, or 'full cabinet' NB-UVB, suitable for larger BSA of vitiligo. While clinical advice to the EAG was that there may be variable access to handheld devices in different NHS trusts, the EAG received clinical expert advice that people with a BSA <10% may still receive NB-UVB therapy. The EAG advisor noted that NB-UVB therapy could be prescribed to any person who has not responded to topical 1st line treatments and wishes to pursue a further active treatment. The EAG also noted that a requirement for >10% BSA was not specified as eligibility criteria for NB-UVB therapy in the BAD guidelines. Overall, the EAG agreed that it was plausible that there will be a group of people who would not choose to receive a topical treatment (at least as a monotherapy) if their vitiligo had a large BSA but did not consider it clear that those with a smaller BSA would not consider NB-UVB.

Source: Company clarification response (question B1)

Table 5: Summary of decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Company rationale if different from the final NICE scope	EAG comment
Population	People aged 12 years and older with NSV with facial involvement	Adults and adolescents from 12 years of age with NSV with facial involvement for whom the disease has not responded to TCS or TCI, or for whom TCS or TCI are contraindicated, not tolerated or otherwise medically inadvisable.	NA	Clinical effectiveness evidence presented by the company was consistent with the NICE scope; i.e. people aged 12 years and older with NSV and facial involvement. However, the economic analysis presented by the company was based on a sub-population of the NICE scope population, limited to people who have previously received treatment (although the choice of the comparator used in the economic analyses was inconsistent with the use of this population). In principle, the EAG did not disagree with the positioning of ruxolitinib as a 2nd line treatment option, though as noted in Key Issue 2, did not consider that the CS was consistent with this population. The EAG was unable to determine whether the clinical effectiveness of ruxolitinib varied according to treatment line. Clinical advice to the EAG was that topical treatments, including ruxolitinib, may not be appropriate for people whose condition covers a large body area. The licence for ruxolitinib limits use of ruxolitinib to be applied to less than 10% BSA. However, the EAG expected that those with a higher overall BSA may still be eligible to

	Final scope issued by NICE	Decision problem addressed in the company submission	Company rationale if different from the final NICE scope	EAG comment
				apply ruxolitinib to some of their vitiligo patches up to this BSA. Clinical advice to the EAG was that those with rapidly progressing vitiligo may be referred for NB-UVB therapy or oral treatment with betamethasone rather than for topical treatments. However, the EAG noted that those with rapidly progressing disease are not precluded from receiving ruxolitinib.
Intervention	Ruxolitinib cream	Ruxolitinib cream	NA	In the trials, ruxolitinib could be used alongside inactive management strategies, such as camouflage make-up, sunscreen and emollients. No active treatments for vitiligo were permitted during the TRuE-V trials. The EAG's clinical advisor stated that they would not consider prescribing ruxolitinib in combination with other active treatments, due to the lack of evidence for the safety of this approach. The licence for ruxolitinib[1] also advises against using ruxolitinib in combination with other topical medicinal products in the same skin areas. However, the EAG considered it plausible that some clinicians may prescribe ruxolitinib in combination with other treatments, including topical treatments used on separate body areas.
Comparator(s)	Established clinical management without ruxolitinib cream	Vehicle cream	To date, established clinical management involved the use of off-label treatments, which consist	As outlined in Key Issue 1, the EAG disagreed with the company's definition of the relevant comparator as vehicle

	Final scope issued by NICE	Decision problem addressed in the company submission	Company rationale if different from the final NICE scope	EAG comment
			of TCS, TCI, phototherapy, laser therapy, topical vitamin D analogues, and a combination of phototherapy with TCI/TCS. Ruxolitinib cream is anticipated to be positioned as a step change option between first and second line for adults and adolescents from 12 years of age with NSV with facial involvement for whom the disease has not responded to TCS, TCI, or for whom TCS or TCI are contraindicated, not tolerated or otherwise medically inadvisable. Therefore, TCS, TCI and phototherapy are not relevant comparators. Given the lack of treatment alternatives in the anticipated positioning, vehicle cream as investigated in the double-blind phase of the TRuE-V trials[7,8] is an appropriate comparator for the appraisal of ruxolitinib cream. Notwithstanding this positioning in the treatment pathway, an ITC FA was conducted to also investigate the feasibility of deriving treatment effect estimates for ruxolitinib cream relative to TCS, TCI and phototherapy. The ITC FA found that there is an insufficient evidence base to robustly compare the	cream, which was not used as a treatment for vitiligo. While in principle the EAG accepted the proposed positioning for ruxolitinib (between 1st and 2nd line), the clinical decision to use ruxolitinib would therefore be a decision between ruxolitinib and existing 2nd line treatments. The EAG therefore concluded that the relevant comparison was between ruxolitinib and existing 2nd line treatments. Several treatments used for vitiligo are not currently available within the NHS, including excimer laser therapy and skin grafting. Established clinical management was considered by the EAG to include those treatments typically used within the NHS in addition to non-active strategies, such as camouflage make-up and sunscreen.

	Final scope issued by NICE	Decision problem addressed in the company submission	Company rationale if different from the final NICE scope	EAG comment
			efficacy of ruxolitinib cream to existing off-label therapies.	
Outcomes	The outcome measures to be considered include: • Re-pigmentation • Maintenance of response • Cessation of spread or stabilisation of vitiligo • Global assessment of vitiligo • Cosmetic acceptability • Adverse effects of treatment • Health related quality of life (HRQoL).	Incyte agrees that the suggested outcomes are appropriate, but notes that stabilisation of vitiligo was not captured in the TRuE-V[7] studies. However, Incyte deems that the endpoint of time to relapse (< F-VASI75) in the long-term treatment extension study (TRuE-V LTE[9]) adequately captures the maintenance of response to treatment.	NA	The company presented evidence for all of the scoped outcomes, thought agreed with the company that evidence for cessation of spread and stabilisation of vitiligo was based on the assessment of relapse rates in the TRuE-V-LTE trial.

	Final scope issued by NICE	Decision problem addressed in the company submission	Company rationale if different from the final NICE scope	EAG comment
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	As per the scope	NA	The company presented an economic analysis that is in keeping with the reference case. The time horizon specified is sufficient but may be considered excessive.
Subgroups	Not included in the draft scope	Due to the anticipated positioning of ruxolitinib cream, the subgroup “prior therapy” is used in the base case, and additional analyses	Vitiligo is more noticeable in people with darker skin tones and associated with higher disease burden, therefore differential cost-effectiveness is expected in this subgroup. A request was made during the decision problem	The company presented evidence according to Fitzpatrick skin type, though noted that the comparison reported was different to that the company stated was requested in the decision problem meeting (the company presented a comparison between Fitzpatrick scale Type I/II and

	Final scope issued by NICE	Decision problem addressed in the company submission	Company rationale if different from the final NICE scope	EAG comment
		are presented using the intention-to-treat (ITT) population and the subgroup "Fitzpatrick Skin Type IV-VI".	meeting that Incyte presents this subgroup analysis.	Type III/IV/V/VI). The EAG was uncertain if this analysis would fully determine whether those with the darkest skin types experience a differential treatment effect. The company suggested that this would be the case.
Special considerations including issues related to equity or equality	Not included in the draft scope	No equality issues are foreseen in terms of providing ruxolitinib cream	Although vitiligo is more noticeable in people with darker skin tones, as noted in the draft scope, and while we expect differential cost-effectiveness in this subgroup due to the different impact of repigmentation on HRQoL, Incyte aims to make ruxolitinib cream available for all patients. Therefore, no equality issues are foreseen in terms of providing ruxolitinib cream to eligible patients, including adults and adolescents from 12 years of age.	The EAG did not identify any equality issues for this appraisal.

Abbreviations: BSA, body surface area; EAG, External Assessment Group; HRQoL, health-related quality of life; NB-UVB, narrowband ultraviolet B therapy; NICE, National Institute for Health and Care Excellence; NSV, non-segmental vitiligo; TCI, topical calcineurin inhibitors; TCS, topical corticosteroids

3. CLINICAL EFFECTIVENESS

3.1. Critique of the methods of review(s)

The company undertook a systematic literature review (SLR) to identify and summarise the comparative efficacy and safety of treatment options (either as independent or as combination therapy) available for people with vitiligo, including ruxolitinib cream. The search strategies, eligibility criteria, screening, data extraction, and quality assessment appeared appropriate. Overall, the EAG found the company's SLR methods to be of reasonable quality and, if followed, would likely have identified all relevant studies for the appraisal. However, the EAG conducted a simple search of the trial registers using terms for 'ruxolitinib' and 'vitiligo' and found six additional records (nine rather than the three reported in the CS). As these were all trials of ruxolitinib, the EAG was unsure why they were not identified by the company in the SLR and presented in the CS.

A summary of the EAG's critique of the methods implemented by the company to identify evidence relevant to the decision problem is presented in Table 6. The company used the results of the SLR to assess the feasibility of an indirect treatment comparison (ITC) and these methods are critiqued in section 3.4.

Table 6: Summary of EAG's critique of the methods implemented by the company to identify evidence relevant to the decision problem

Systematic review step	Section of CS in which methods are reported	EAG assessment of robustness of methods
Searches	Appendix D	The search strategies were well structured and executed with a good range of sources. Terms for vitiligo were appropriately combined with terms for ruxolitinib and comparators. To this was added a broad filter for clinical trials and prospective studies. Case reports and conference abstracts were excluded and results were limited to English language only. The company carried out clinical trials searches in WHO ICTRP and in clinicaltrials.gov, the strategies used were not described. The EAG carried out trial searches in the same two sources using a simple strategy (vitiligo AND (ruxolitinib OR opzelura)) and found nine trial records in contrast to the three trials in the CS (plus a further trial mentioned in clarification). It was not possible for the EAG to appraise the additional trials within the timeframe of the appraisal, although the trials appeared to include two completed trials of ruxolitinib for the treatment of vitiligo that included clinical efficacy outcomes relevant to this appraisal.

Systematic review step	Section of CS in which methods are reported	EAG assessment of robustness of methods
Inclusion criteria	Appendix D, Table 6	The eligibility criteria used in the SLR was wider than that of the Final scope issued by NICE. For example, the population included adolescents and adults diagnosed with "any type" of vitiligo, rather than limiting to people with NSV. While EAG did not consider this was a risk that relevant studies had been missed, it led to the SLR containing studies with limited applicability to the decision problem.
Screening	Appendix, D1.1.3	The EAG considered the methods for screening to be adequate.
Data extraction	Appendix, D1.1.4 and D1.1.6	The EAG was satisfied with the data extraction process.
Tool for quality assessment of included study or studies	Appendix, D1.1.9	The EAG noted that the quality assessment presented in Document B used the CRD's "minimum criteria for assessment of risk of bias in RCTs". No additional sources of potential bias were considered, and the tool was used inappropriately to assess the single arm component (Cohort B) of the TRuE-V-LTE trial. The EAG identified additional risks of bias in the key trials for ruxolitinib that were not identified by the company. All RCTs identified by the company in their SLR were assessed using the Cochrane Risk of Bias assessment (RoB 2) tool[10] and used suitable tools for the non-randomised and single-arm trials.
Evidence synthesis	SLR[11]	As noted above, the inclusion criteria led to the SLR including 253 studies, a proportion of which have limited applicability to the decision problem. The company presented outcomes from the included studies in tables that were ordered by study design and by treatment group. No meta-analysis or narrative synthesis was undertaken. The EAG accepted that a proportion of the studies had limited applicability, however evidence synthesis focusing on studies that were closely related to the decision would have supported decision-making. This could have been a narrative synthesis of studies in the NSV population where treatments relevant to the decision problem such as TCS, TCI, and NB-UVB, alone or in combination, were compared to each other or to a placebo treatment such as vehicle cream. This synthesis would have contextualised the evidence landscape surrounding the decision problem and have provided further clarity about the feasibility of an ITC.

Abbreviations: CS, Company submission; EAG, External Assessment Group; ITC, indirect treatment comparison; NSV, nonsegmental vitiligo; SLR, systematic literature review; TCI, topical calcineurin inhibitors; TCS, topical corticosteroids.

3.2. Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these)

3.2.1. Studies included in the clinical effectiveness review

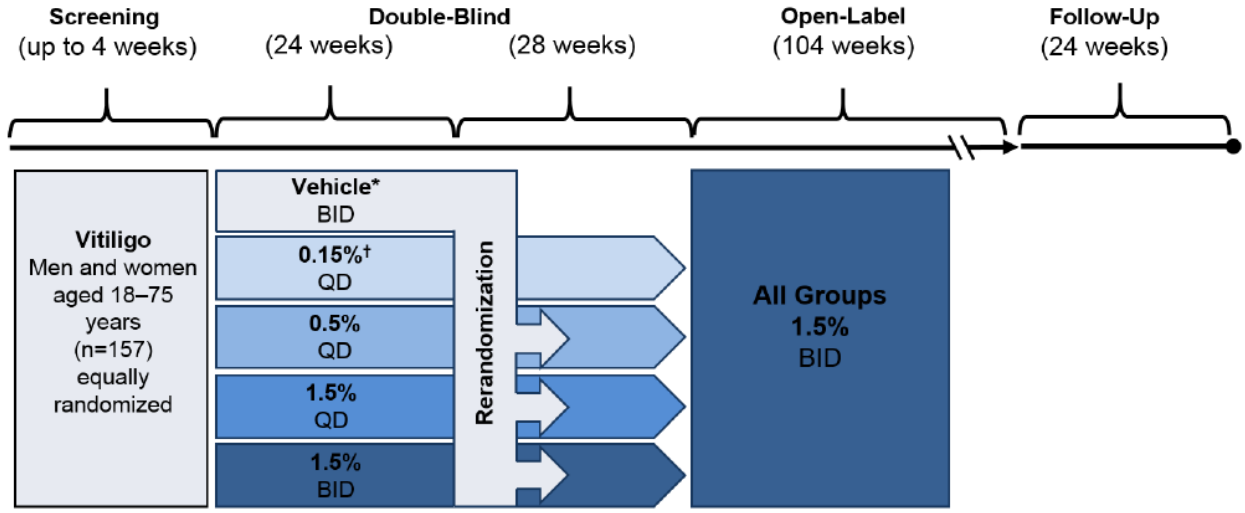
The CS described two trials: TRuE-V1[7] and TRuE-V2[8] (Table 7). These were two 'identically designed' international phase 3, double-blind, randomised controlled trials (RCTs). In each trial,

treatment with ruxolitinib was compared to vehicle cream (i.e. a placebo intervention). The trials each included a 24-week double-blind phase, after which point treatment was unblinded and participants who received vehicle cream could choose to switch to ruxolitinib up until the end of the trial (the open-label extension [OLE]; final follow-up 52 weeks). At the end of the OLE, those participants who had complied with treatment, completed sufficient outcome measures and showed no safety concerns were eligible to participate in a further trial extension (TRuE-V LTE[9]). In this trial, those who responded to treatment in the earlier trial phases entered a double-blind RCT comparing either continuation with ruxolitinib (long-term treatment) or switching to vehicle cream (withdrawal). Those who were allocated to vehicle cream in the LTE could restart ruxolitinib following relapse, and therefore the trial also provided evidence on the management of relapse. Those who did not respond to treatment during the earlier trial phases (as defined by less than 90% facial repigmentation) entered an open-label, single arm evaluation of continued ruxolitinib.

The EAG was aware that a phase II trial of ruxolitinib had also been conducted. This was identified in the company's SLR but no details of the trial or its findings were reported in the CS. This was queried by the EAG at clarification (question A15), and in response the company provided the CSR for the trial[12], though on examination the version provided by the company did not contain full data for the trial (notably, missing safety data for the double-blind phase of the trial). The company did not provide a rationale for why evidence from this trial had not been presented in the CS.

The Phase II trial was a randomised, double-blind, vehicle-controlled, dose-finding study in adult participants with vitiligo (≥0.5% of facial BSA and ≥3% of non-facial BSA). Four doses of ruxolitinib were evaluated: 0.15% QD, 0.5% QD, 1.5% QD, 1.5% BID (the latter, highest dose being the dose evaluated in the TRuE-V trials). After 24-weeks of treatment, those in the vehicle arm and those on the lowest dose (0.15% QD) who had not achieved a response were re-randomised to one of the higher dose ruxolitinib arms (still blinded). After 52 weeks from baseline, participants with no safety concerns, no clinically significant changes in laboratory parameters, and had completed sufficient assessments were invited to participate in a further 104 week open-label extension period (study design shown in Figure 1).

Figure 1: Phase II trial study design



* Rerandomization to 0.5% QD, 1.5% QD, or 1.5% BID at Week 24 for vehicle group.
† Rerandomization to 0.5% QD, 1.5% QD, or 1.5% BID if <25% improvement in F-VASI at Week 24.

BID = twice daily; F-VASI = facial Vitiligo Area Scoring Index; QD = once daily.

Source: INCB 18424-211 CSR[12]

The EAG considered that evidence from the Phase II trial should have been provided in the CS for this appraisal. Within the double-blind 52-week period, the trial could provide information about the safety of ruxolitinib (i.e. before participants with 'safety concerns' from ruxolitinib were excluded from the trial). Moreover, clinical outcomes for the 1.5% BID arm could be compared with the findings of the Phase III trials, and a comparison between dose arms could provide information about the possibility of a dose response (this issue is of relevance to this appraisal, see Section 3.2.2.3and 4.2.4). Within the timeframe of this appraisal, the EAG was unable to fully appraise the evidence from this trial.

Finally, shortly before submission of this report, the EAG identified two other Phase II trials of ruxolitinib that were not described in the CS. These were:

- TRuE-V MOA - NCT04896385 A Study to Evaluate the Mechanism of Action of Ruxolitinib Cream in Subjects With Vitiligo (TRuE-V MOA) [Completed]. Phase 2.[13]
- NCT02809976 - Topical Ruxolitinib for the Treatment of Vitiligo [Completed]. Phase 2.[14]

At the same time, the EAG also identified three additional ongoing trials not described in the CS. These were:

- NCT05750823 - A Study to Assess the Safety and Efficacy of Ruxolitinib Cream in Participants with Genital Vitiligo [Recruiting]. Phase 2.[15]
- NCT05247489 - A Study to Evaluate the Safety and Efficacy of Ruxolitinib Cream With Phototherapy in Participants With Vitiligo [Active, not recruiting].Phase 2.[16]
- NCT05872477 - Promoting Repigmentation After Epidermal Cell Suspension Grafting and preVENTing the Loss of Melanocytes Using Topical Ruxolitinib for Vitiligo in Resistant Areas (PREVENT) [Not yet recruiting]. Phase 2.[17]

The EAG was unsure why these studies were not identified by the company's SLR or discussed in the CS. The EAG was unable to fully appraise these trials during its appraisal, however identified that the first completed trial (TRuE-V MOA) was a randomised, double-blind, vehicle-controlled trial with an open-label extension that assessed safety and efficacy outcomes and may therefore have been of relevance to this appraisal. The second completed trial was potentially less relevant for consideration, as this was a small (N=11), single-arm trial.

Completed trials of ruxolitinib are shown in Table 7. Trials shown in grey are those for which the company did not provide clinical effectiveness and safety evidence in the CS.

Table 7: Completed clinical trials of ruxolitinib for the treatment for vitiligo

Study name and acronym	Study design	Population	Intervention	Comparator	Study type
TRuE-V1[18]	Double-blind RCT	Adolescents and adults aged ≥ 12 years with NSV affecting the face (≥ 0.5% BSA on the face, ≥ 0.5 F-VASI) and ≥ 3% BSA on non-facial areas, ≥ 3 T-VASI, and total body vitiligo area (facial and non-facial) not exceeding 10% BSA N=330	Ruxolitinib	Vehicle cream	Clinical efficacy and safety
TRuE-V2[19]	Double-blind RCT	As TRUE-V1 N=344	Ruxolitinib	Vehicle cream	Clinical efficacy and safety
TRuE-V LTE[9]	Double-blind RCT [Cohort A – those who responded to ruxolitinib during the previous trials] followed by an open-label extension in those who relapsed Open-label single-arm trial [Cohort B – those who did not respond to ruxolitinib during the previous trials]	Participants from TRuE-V1 and TRuE-V2 who had complied with treatment up to the final follow-up and showed no safety concerns Cohort A N = 116 Cohort B N = 342	Continuation with ruxolitinib	Discontinuation of ruxolitinib (to vehicle cream); Re-initiation of ruxolitinib following relapse	Clinical efficacy and safety
INCB 18424-211[12]	Double-blind RCT	Adults aged 18-75 years with vitiligo	Alternative doses of ruxolitinib	Vehicle cream	Dose-finding, safety

		N=157		Dose comparison	
TRUE-V MOA[13]	Double-blind RCT	Adults with NSV affecting the face (≥ 0.5% BSA on the face, ≥ 0.5 F-VASI) and ≥ 3% BSA on nonfacial areas, ≥ 3 T-VASI; total body vitiligo area (facial and nonfacial) not exceeding 50% BSA. N=60	Ruxolitinib	Vehicle cream	Clinical efficacy and safety
NCT02809976[14]	Single-arm trial	Adults with vitiligo covering at least 1% of total BSA N=11	Ruxolitinib	None	Clinical efficacy

Abbreviations: RCT, randomised controlled trial

Note: Trials that are greyed out did not form part of the CS and were not appraised by the EAG during this appraisal.

3.2.2. Description and critique of the design of the studies

3.2.2.1. Design of the studies

The two main trials for ruxolitinib presented in the CS, TRuE-V1 and TRuE-V2, used the same design: these were double-blind, randomised, placebo (vehicle cream)-controlled trials with 2:1 randomisation, followed by a single-arm, open-label extension. Randomisation was stratified by geographic region (North America vs Europe) and Fitzpatrick skin type. More than two thirds of trial centres were based in North America. No centres were based in the UK but the EAG was unaware of any rationale to suggest that this would limit the generalisability of the trial data. The EAG agreed with the company's rationale for pooling the two trials: while minor variations in participant demographics and clinical outcomes were noted between the two trials, the EAG agreed that the trials were of the same design and that pooling would provide a better representation of the clinical effectiveness of ruxolitinib.

As noted in Key Issue 1, the EAG did not consider the choice of vehicle cream as the trial comparator to be informative for determining the appropriate positioning of ruxolitinib in the treatment pathway, or for informing cost effectiveness estimates. However, the EAG considered the design to be acceptable for determining whether ruxolitinib was clinically effective as compared to no treatment.

The TRuE-V-LTE trial included two thirds (68.0%) of participants from the TRUE-V1 and TRuE-V2 trials. These were participants who completed the previous trials with good compliance and who tolerated ruxolitinib without safety concerns. The trial split participants into two cohorts depending on their response to ruxolitinib in the previous trial phases: Cohort A was comprised of participants who had an excellent response to ruxolitinib treatment by 52 weeks (as defined by 90% repigmentation, F-VASI90) and Cohort B was comprised of participants who had not shown this level of response by week 52. As participants included those who had been randomised to vehicle cream during the initial 24-weeks of the trials, the assessment of whether participants had responded to treatment was based on a timeline ranging between 28- and 52-weeks. Those in Cohort A (responders) were randomised to either continuation with ruxolitinib or discontinuation to vehicle cream (double-blind). The findings of this analysis were useful for assessing maintenance of response in those either continuing or withdrawing from treatment. Those in Cohort B (non-response or response <F-VASI90) all continued to receive open-label ruxolitinib. The findings of this analysis were useful for assessing clinical outcomes with longer treatment duration. The EAG considered the findings of the TRuE-LTE trial to provide an insight

into longer term outcomes, including whether the effect of ruxolitinib would be maintained over time (with or without continued treatment). However, as the trial was limited to a sub-sample of the original trials, the selection of which may be open to selection bias, the EAG considered that the findings should be interpreted with caution. Moreover, the EAG noted that the threshold used to determine response (F-VASI90) was higher that the threshold for a response used by the company elsewhere in the submission (F-VASI70) and supported by clinical advice to the EAG. The findings of the TRuE-V-LTE therefore had limitations in generalisability that need to be considered when interpreting the findings.

The double-blind phase of the trials had a follow-up of 24 weeks while the open-label phase was 28 weeks, thus resulting in a combined follow-up of 52 weeks. Clinical advice to the EAG suggested that the mechanism of ruxolitinib would result in a gradual response over time, which was supported by the clinical effectiveness data. Treatment response was shown to increase in a minority of participants up to the 52-week follow-up, suggesting that for the vast majority of participants, the trial follow-up was sufficient for assessing treatment response. Further follow-up of people continuing on ruxolitinib was available in the TRuE-V-LTE trial (up to 103 weeks).

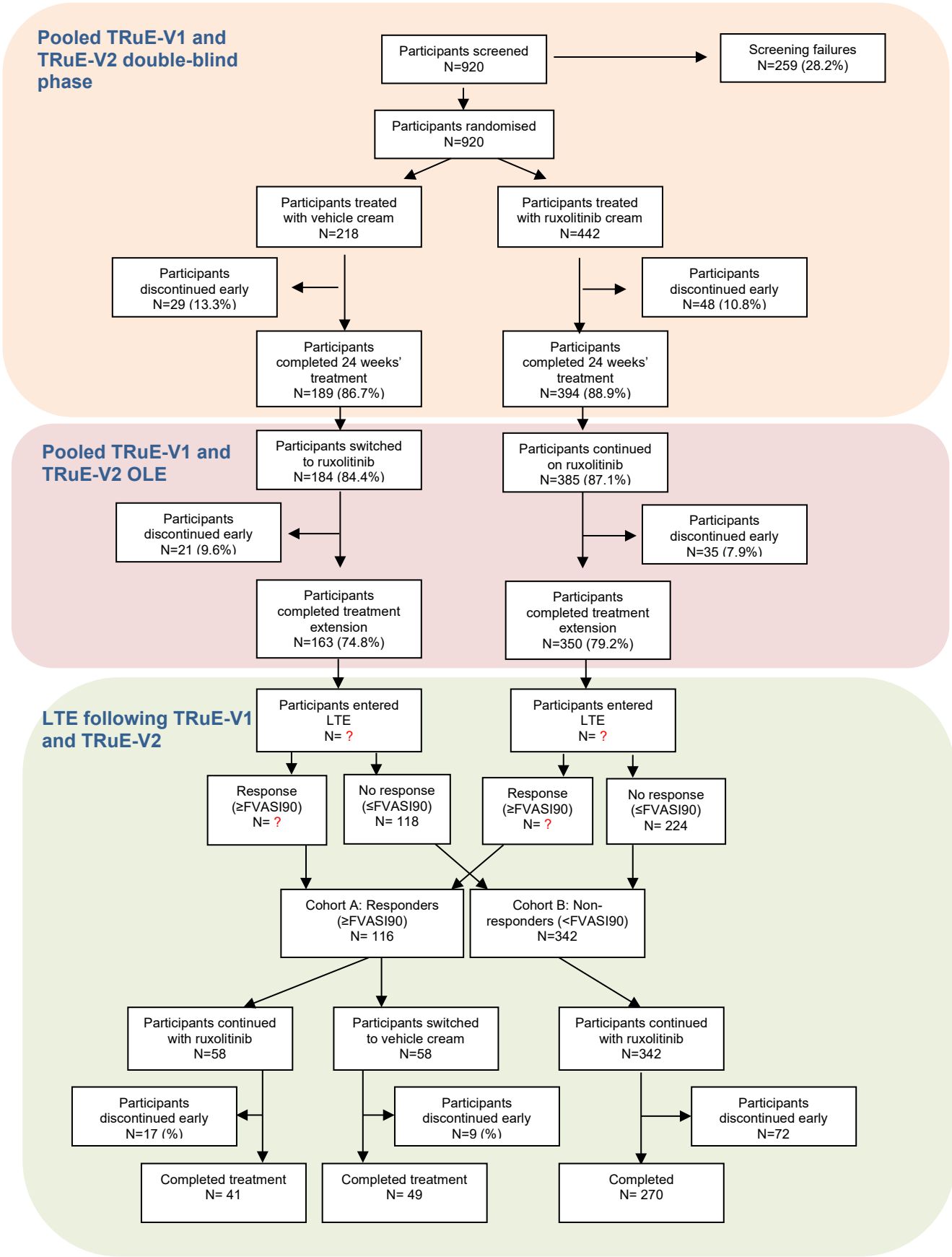
However, the EAG was less clear to what extent the length of follow-up was appropriate for determining maintenance of the treatment response. Data from the TRuE-V-LTE trial suggested that further follow-up was needed to determine the typical duration of response. The TRuE-V-LTE trial reports treatment efficacy after one round of re-treatment with ruxolitinib in those with a high level of response (F-VASI90). However, a limitation of the trials is that the efficacy of retreatment for those with a lower prior response was not captured, nor was the efficacy of multiple rounds of re-treatment. This issue is discussed further in Sections 4 and 6.

The EAG considered that the psychological impacts of change in vitiligo outcomes may take longer to demonstrate and may not be evident during the double-blind phase of the trials. However, the EAG considered that the 52-week follow-up and beyond would be a reasonable timeframe for evaluating these outcomes.

Finally, while the EAG considered that the trials were sufficiently long to capture any immediate adverse effects, the SmPC[1] for ruxolitinib noted that the trials may not be long enough to assess whether ruxolitinib was associated with any meaningful long-term risks. Specifically, as discussed in Section 3.2.3.1, the trials were unlikely to capture the risk of nonmelanoma skin cancer.

Participant flow across the different phases of the trials and their extensions was complicated and the EAG found conflicting numbers for each group and phase in the CS and trial CSRs. The EAG suspected that this was due to the data being reported in separate sections that refer to different analysis sets. The EAG has included an overview of the participant flow through the different study phases in Figure 2, though due to the reason above, these numbers may differ from those reported in places in the CS. The EAG was unable to identify participant numbers for some stages of the participant flow. The company stated that data was missing from the analyses due to missing assessments during the COVID-19 pandemic and the exclusion of data from one of the trial sites (site 710). The number of participants discontinuing from the TRuE-V1 and TRuE-V2 trials was limited, with fewer than 10% of participants discontinuing in each arm. However, there were moderate levels of drop out during the TRuE-V-LTE, and across the trial phases this resulted in an approximate 20% loss of participants treated with ruxolitinib in the TruE-V-LTE endpoints. The EAG was unclear to what extent the missing data was due to drop-out from the trial or whether a number of participants did not meet the company's criteria for entry (i.e. no safety concerns with continuing ruxolitinib). In general, across the trial phases, the most common reasons for drop-out were those that could plausibly relate to efficacy or safety (e.g. withdrawal by participant, loss to follow-up). Given the magnitude of missing data and the potential for missing data to not be missing at random, the EAG considered that a 20% discontinuation rate could meaningfully affect treatment outcomes in the TRuE-V-LTE, and that appropriate missing data analysis would be influential. However, while imputation of missing data was conducted for the TRuE-V1 and TRuE-V2 trials (where missing data was minimal), this was not conducted for the TRuE-LTE (based on evidence in the CS). The EAG therefore considered that efficacy data from 52 weeks onwards in the submitted evidence base was at a high risk of attrition bias (see Section 3.2.2.6).

Figure 2: Participant flow in the pooled TRuE-V1 and TRuE-V2 trials, including the OLE and LTE



Note: aside from screening failures, where % represents the proportion of those screened who did not meet trial eligibility criteria, all other %s are calculated using the number of participants who received treatment during the double-blind period as the denominator.

Abbreviations: LTE, long-term extension; OLE, open-label extension

3.2.2.2. Population

The population in the NICE final scope was people aged 12 years and older with NSV with facial involvement. The TRuE-V1[18] and TRuE-V2[19] trials were based in 45 and 49 study centres respectively, and these centres were located in North America and Europe. No centres were located in the UK but the EAG understand the vitiligo care received in the site locations to be generalisable to the UK.

Trial eligibility criteria

Eligibility criteria for the TRuE-V1[18] and TRuE-V2[19] trials were provided in the CS (Document B, Table 6). The MHRA granted marketing authorisation for ruxolitinib[1] was broadly consistent with the eligibility criteria of the TRuE-V1 and TRuE-V2 trials. The therapeutic indication for ruxolitinib was treatment of NSV with facial involvement, however the license does not restrict to use on the face and it can be applied to any depigmented skin areas. The company have positioned ruxolitinib as a 2nd line treatment option, after TCS and TCIs, and if approved in this position, the population treated in NHS practice would likely be a subset of the population recruited to TRuE-V1 and TRuE-V2 trials. This was further discussed in Section 2.4 and Key Issue 2.

Prior to randomisation, 920 participants were screened for inclusion in the trial and 259 (28.2%) were deemed to be 'screening failures'. At the clarification stage (question A12), the EAG requested comment from the company on what appeared to be high numbers of screening failures and the reasons for this. The company stated that the specific criteria not met during screening were not collected. The EAG's clinical expert considered that the trial participants were nevertheless representative of the target population.

Participants who were enrolled and receiving treatment in either TRuE-V1[18] or TRuE-V2[19], were currently tolerating ruxolitinib cream, and presented no safety concerns for investigators, were invited to join the TRuE-V LTE[9] treatment extension trial. As such, this trial did not represent the whole target population of people eligible to receive ruxolitinib, but instead represented a subset of people who tolerated treatment. There were two components within TRuE-V LTE based on a

person's response to treatment in the TRuE-V1/TRuE-V2. People who had F-VASI90 at the end of TRuE-V1/TRuE-V2 entered a comparative trial and were randomised to either ruxolitinib or vehicle cream. People who did not have F-VASI90 joined a single arm trial using ruxolitinib. The EAG noted that in other places in this submission F-VASI75 is defined as a clinically meaningful response but that is not considered sufficient to enter the TRuE-V LTE RCT component.

Baseline characteristics

The demographic characteristics and baseline disease characteristics of the ITT populations from the TRuE-V1[18] and TRuE-V2[19] trials are reported in Table 7 and Table 8 of the CS (Doc B). Full baseline characteristics were not presented solely for the previously treated subgroup.

The treatment groups were well-balanced in demographic characteristics. Five hundred and fifty-two (81.9%) were White, 32 (4.7%) Black/African American, 28 (4.2%) Asian, 37 (5.5%) other, and 25 (3.7%) not reported. The race of participants was broadly representative of the UK 2021 Census data that reported that 82% of people in England and Wales were White and 18% belonged to a Black, Asian, mixed or other ethnic group[20]. Subgroup analysis presented in Figure 18 of the CS (Doc B), reported a similar response rate in the proportion of participants reaching F-VASI75 at week 24 across the race categories.

Participants' skin types were assessed using the Fitzpatrick skin phototypes (FSP) scale that classified skin from types I to VI. The original FST classifications included skin types I through IV; skin types V and VI were later added which correspond to people of Asian, Indian, and African origin. Most of the participants in the trial had Fitzpatrick Skin Type II, III and IV (88.9%). It was unclear if the skin types of the participants were representative of the population with vitiligo in England and Wales, but as noted, the race of the participants was broadly in line with the 2021 census data.

The baseline disease characteristics were also well-balanced between treatment groups. The mean time since diagnosis was 14.79 years with a median of 11.97 years since diagnosis. The EAG's clinical expert considered that this was consistent with more long-standing disease and noted that people with long-standing vitiligo may be less responsive to treatment. The company did not present subgroup data to compare treatment response according to time since diagnosis, and the EAG was therefore unclear to what extent this would be a treatment effect modifier.

As noted in Section 2.4, the company proposed positioning ruxolitinib as a new line of therapy in between the current 1st and 2nd line therapy in the BAD guidelines[21]. This would position it as a treatment for people whose condition had not responded to TCS and/or TCI, or for whom TCS or TCI are contraindicated. This would be prior to use of NB-UVB with or without TCS or TCI or oral betamethasone for those with rapidly progressing disease. Sixty-one per cent of participants in the TRuE-V1/TRuE-V2 trials had received prior therapy for vitiligo (Table 8, Doc B). Similar proportions in the trial had previously used TCS (28.0%), TCIs(31.8%), and NB-UVB (31.9%), to treat their vitiligo. Based on the evidence presented by the EAG, it was not possible to determine the proportion of participants in the trials for whom the disease had not responded to TCS and/or TCI, or for whom TCS or TCI were contraindicated, not tolerated or otherwise medically inadvisable. It was also not possible to determine the overlap between the number of participants who had previously received each previous treatment. However, a proportion of the trial participants had not received either TCS or TCSs at baseline, while nearly a third of participants had received NB-UVB, a later line of treatment. It was unclear how generalisable the full trial population was to the proposed 2nd line population for ruxolitinib.

The demographic characteristics and baseline disease characteristics of the participants in the TRuE-V LTE trial were reported in Table 9 and Table 10 of the CS (Doc B). The treatment groups in Cohort A were well balanced for baseline demographic and disease characteristics. All participants entering Cohort B were treated with ruxolitinib and were presented in groups based on their treatment arm in the TRuE-V1 and TRuE-V2 trials.

The participants entering the TRuE-V LTE (Cohort A and Cohort B) trial were a subset of those recruited to TRuE-V1/TRuE-V2, who tolerated treatment and wished to continue in the trial. The participants entering Cohort A and Cohort B had similar demographics and baseline characteristics to those recruited to TRuE-V1/TRuE-V2. However, the EAG noted that a higher proportion of participants in Cohort A had received prior therapy for vitiligo than the ITT population in TRuE-V1/TRuE-V2 (71.6% compared to 61.0%).

3.2.2.3. Intervention

Participants randomised to ruxolitinib applied the treatment twice daily for 24 weeks to all vitiligo areas on the face and body. Consistent with the product licence, the recommended dose was a thin layer of cream applied twice daily to the depigmented skin areas up to a maximum of 10% of BSA, with a minimum of 8 hours between two applications[1]. Ten per cent BSA represents an area as large as 10 times the palm of one hand with the 5 fingers. In the trials, participants were

given one 60-gram tube of ruxolitinib each week, equivalent to up to 240 grams over a four-week period. This is inconsistent with the product licence for ruxolitinib, which specified that no more than two tubes of 100 grams a month should be used.

A summary of exposure was presented in Table 37 of the EMA SmPC reported (provided in Appendix C of the CS), which has been adapted below in Table 8. The median weight of ruxolitinib applied in the trials was 4.07 grams per day but the mean (SD) dose was substantially higher at 7.36 (25.2) grams per day. Also, the EAG noted that the maximum dose applied was 237.1 grams per day. Therefore, at least one participant was applying substantially more ruxolitinib each day than the licence indicates. Based on the data provided by the company, the EAG was unable to determine how many trial participants used more than the licenced dose of ruxolitinib. The EAG was unclear how higher doses of ruxolitinib would affect clinical outcomes in the trial. The Phase II trial of ruxolitinib (see Section 3.2.2) included a dose comparison and showed that increased efficacy is possible with higher doses of ruxolitinib, but the highest dose used in the trial was the licensed dose and so efficacy evidence is not available for a higher dose. The safety implications of higher ruxolitinib doses were also unclear. The dose of ruxolitinib and the implications of assumptions around dosing is further discussed in relation to the economic model in section 4.2.4 and in Key Issue 4.

The licence for ruxolitinib specified that it should be used cutaneously only and that people should avoid washing treated skin for at least two hours after application.[1] Other topical medicinal products used to treat other conditions on the same skin areas should be applied with a minimum of two hours after the application of ruxolitinib. This was also applicable to the use of sunscreen or emollients, and in the trial these were required to be removed from the skin prior to applying ruxolitinib. In the trials, this restriction also applied to the use of camouflage make up. The EAG considered that these restrictions were sensible but may nevertheless be challenging to adhere to in practice. People with vitiligo are encouraged to maintain consistent use of sunscreen to protect depigmented skin, and people may also use camouflage make-up to reduce the impact of their condition. The EAG considered it plausible that in practice, adherence to these restrictions may be challenging around daily activities. This may reduce the effectiveness of ruxolitinib in practice. Clinical advice to the EAG was that the application of topical treatments is burdensome for people with vitiligo, and therefore the use of ruxolitinib with these restrictions may be equally or more burdensome.

Compliance of > 80% of the drug applications over the double-blind period for both the ruxolitinib and vehicle cream treatment arms in the TRuE-V studies was 98.1% (Table 3.1.2.1 in the TRuE-V1 CSR and TRuE-V2 CSR). However, participants (n = 13) from study site 710 in TRuE-V2 were excluded from the study due to non-compliance with the protocol and concerns with data quality. The EAG requested clarification (question A7) on how these participants were identified and the company noted that participants were not excluded from site 710, but that data from all participants at site 710 was excluded. The company did not offer any specific detail of the non-compliance with the protocol or why there were concerns with data quality. In the EMA SmPC report, the authors reported that the decision to exclude the data was due to "one critical finding (informed consent) and two major findings (source documents and organisation and personnel)" (p.103). Given the low numbers of people excluded from the trials, the EAG did not consider that the exclusion would meaningfully affect clinical outcomes, but still considered this to be an uncertainty in the appraisal.

Table 8. Summary of exposure in TRuE-V1/TRuE-V2 (adapted from Table 37, CS Appendix C [SmPC report])

Variable	Vehicle cream BID (N=224)	Ruxolitinib 1.5% cream BID (N=449)	Total (N=673)
Duration of treatment (days)			
Mean (SD)	156.8 (38.9)	158.9 (35.0)	158.2 (36.3)
Median	168.0	168.0	168.0
Min, max	1.0, 248.0	1.0, 237.0	1.0, 248.0
Average weight of medication applied (g)			
Mean (SD)	7.12 (22.96)	7.36 (25.23)	7.28 (24.48)
Median	3.81	4.07	4.03
Min, max	0.3, 236.3	0.4, 237.1	0.3, 237.1

Abbreviations: BID, twice daily; SD, standard deviation

In section 6.6 and 6.7 of the TRuE-V1 and TRuE-V2 protocols, the company described the treatments, vaccinations, and devices allowed or disallowed before, during, and/or after study treatment. Participants were permitted to use bland emollients, camouflage makeup, and a mineral-based sunscreen at least 2 hours after study drug application. However, participants should not use any other treatments for vitiligo at any time during the study. This included corticosteroids (topical, systemic, or oral), vitamin D derivatives, calcineurin inhibitors, laser or surgical treatments, NB-UVB, or other procedures. In addition, skin bleaching treatments, depigmenting agents, biological therapies, immunosuppressant agents, and live or live-attenuated vaccination were not permitted.

In section B.3.5.1 of the CS, the company stated that people receiving ruxolitinib and vehicle cream were assumed to use permitted concomitant therapies including vitamin D supplements, camouflage, fixing powder and sunscreen. The amount used for these therapies was not collected. In addition, a summary of concomitant medications used in the double-blind period was presented in the TRuE-V study CSRs (Table 1.4.3.1) as noted by the company at clarification (question A11). Concomitant medications were received by 73.5% of participants in the ruxolitinib arm, and 72.8% of participants in the vehicle cream arm. Table 1.4.3.1 of the trial CSRs provided details of the number of participants who received each concomitant medication, but the data presented were not sufficient to determine why they received the medication, the formulation of medication (oral/topical/inhaled), the dose, or how often it was used.

3.2.2.4. Comparator

Participants randomised to the control arm applied vehicle cream twice daily for 24 weeks to all vitiligo areas on the face and body. The number of tubes given to participants and the guidance for application was consistent with the ruxolitinib arm (Section 3.2.2.3).

A summary of exposure was presented in the EMA SmPC report Table 37 (CS Appendix C) and has been adapted as shown in Table 8. As with the ruxolitinib arm, the median weight of vehicle cream applied was substantially lower than the mean weight of vehicle cream applied. Also consistent with the ruxolitinib arm, the maximum dose applied far exceeded the intended dose (236.3 grams per day).

A discussion of background treatments received in the control arm can be found above in the Intervention section (3.2.2.3).

3.2.2.5. Outcomes

The outcomes reported in the CS or accessible to the EAG during the timeframe of the appraisal are shown in Table 9. The outcome categories shown correspond to the outcomes specified in the NICE scope for this appraisal.

Response according to the VASI measures, facial or total, comprised the majority of the evidence base for ruxolitinib. The VASI measures consider both BSA and level of pigmentation of vitiligo patches, and so could be considered a composite outcome of these characteristics, both of which are important to people with vitiligo[22]. The company reported response according to various thresholds of change in F-VASI and T-VASI. Research indicates that the level of response considered by people with vitiligo to be meaningful varies across the population.

However, clinical advice and published research suggests that a 75% threshold is considered to be meaningful by most people with vitiligo (i.e. F-VASI75 and T-VASI75). Notably, the company did not account for multiple comparisons in the trial (see Section 3.2.2.6).

As noted previously, clinical advice to the EAG was that VASI assessments of vitiligo are a highly accurate measure vitiligo but are typically not used in practice due to the time needed to perform the assessment. This means that while these outcomes in the trial would be an accurate measure of change in vitiligo lesions, there may be some generalisability issues when interpreting the data (for example, clinical decisions on the basis of response may use alternative criteria in clinical practice).

HRQoL was assessed using three disease-specific instruments, though only data for the VitiQoL was assessed in detail by the EAG during the appraisal. In the CS, the company stated that no difference in HRQoL was found between arms on the DLQI and CDLQI, which are dermatology HRQoL measures, though the data was not presented. The VitiQoL measure was developed to measure the impact of vitiligo on quality of life, including how vitiligo has impacted people's ability to function, their relationships, physical health and emotional wellbeing. The measure has a moderate to poor association with self-reported vitiligo severity[23] and the EAG was unable to find a validated minimally clinical important difference (MCID) threshold, or evidence that the measure was responsive to change. The EAG was unclear why the company had selected not to incorporate a generic measure of HRQoL in its trials, such as the EQ-5D, particularly given the lack of psychometric validation of the VitiQoL. The items in the VitiQoL did not appear to assess additional potential impacts of vitiligo beyond those included in generic HRQoL instruments and this would have reduced uncertainty in the HRQoL effects of treatment.

Table 9: Clinical outcomes for ruxolitinib appraised by the EAG

	Pooled TRuE-V1 and TRuE-V2 **Double-blind phase (24-weeks)**	**Pooled TRuE-V1 and TRuE-V2** **Open-label phase (24 – 52-weeks)**	**TRuE-V-LTE Cohort A (Responders; ≥F-VASI90)** **Double-blind (52 – 103 weeks)**	**TRuE-V-LTE Cohort B (Non-responders; <F-VASI90)** **Open-label (52 – 103 weeks)**
Re-pigmentation	Facial and bodily vitiligo as assessed using F-VASI, F-BSA and T-VASI. Clinician- and patient-rated	Facial and bodily vitiligo as assessed using F-VASI, F-BSA and T-VASI. Clinician- and patient-rated	Facial vitiligo as assessed using F-VASI	Facial vitiligo as assessed using F-VASI

	Pooled TRuE-V1 and TRuE-V2 **Double-blind phase (24-weeks)**	**Pooled TRuE-V1 and TRuE-V2** **Open-label phase (24 – 52-weeks)**	**TRuE-V-LTE Cohort A (Responders; ≥F-VASI90)** **Double-blind (52 – 103 weeks)**	**TRuE-V-LTE Cohort B (Non-responders; <F-VASI90)** **Open-label (52 – 103 weeks)**
	change in facial and total vitiligo.	change in facial and total vitiligo.		
Maintenance of response	Change in F-VASI response	Change in F-VASI response	Relapse in F-VASI	Change in F-VASI response
Cessation of spread or stabilisation of vitiligo	Facial and bodily vitiligo as assessed using F-VASI, F-BSA and T-VASI. Clinician- and patient-rated change in facial and total vitiligo.	Facial and bodily vitiligo as assessed using F-VASI, F-BSA and T-VASI. Clinician- and patient-rated change in facial and total vitiligo.	-	-
Global assessment of vitiligo	T-VASI Clinician- and patient-rated change in total vitiligo	T-VASI Clinician- and patient-rated change in total vitiligo	-	-
Cosmetic acceptability	VNS	VNS	-	-
Adverse effects of treatment	Treatment-emergent AEs	Treatment-emergent AEs	Treatment-emergent AEs	Treatment-emergent AEs
Health-related quality of life	VitiQoL (separate for each trial) HADS	VitiQoL (separate for each trial) HADS	-	-

Abbreviations: AE, adverse events; F-BSA, facial body surface area; F-VASI, Facial Vitiligo Area Scoring Index; HADS, Hospital Anxiety and Depression Scale; T-VASI, Total Vitiligo Area Scoring Index; VitiQoL, Vitiligo-specific quality-of-life instrument; VNS, Vitiligo Noticeability Scale.

The company explored various analyses to accounting for missing data in the TRuE-V1 and TRuE-V2 trials. There was minimal missing data in these trials and the different approaches to analysis did not have a material impact on the results. However, as noted in Section 3.2.2.1, there was high rate of missing data in the TRuE-V-LTE phase of the trial, with a third of people involved in the earlier trial phases not enrolled. The company did not appear to employ the same level of investigation of the effect of missing data in this trial, and the EAG was concerned that missing data could not be determined to be missing at random. As a consequence, the EAG had some concerns about the validity of the results from the TRuE-V-LTE trial (see Section 3.2.2.6).

3.2.2.6. Critical appraisal of the design of the studies

The company assessment of the quality of the TRuE-V1 and TRuE-V2 trials was reported in Appendix D of the CS, Table 9. The assessment utilised version 2 of the Cochrane risk-of-bias tool for randomised trials (RoB 2)[10]. The company concluded that both trials were at a low risk of bias in each domain assessed, but the company did not offer specific reasoning for the judgements.

A further quality appraisal for the TRuE-V1 and TRuE-V2 trials and the TRuE-V-LTE trial was presented in Table 5 (Section B.2.5) of Document B. The assessment for the TRuE-V1 and -V2 trials is presented in the same column, due to the comparability of the trial methods. This assessment was conducted using the "minimum criteria for assessment of risk of bias in RCTs" set out in CRD's guidance for undertaking reviews in health care[24]. The company provided specific reasoning linked to the rating for each domain of the assessment, though no overall risk of bias judgement was made. The company assessment did not appear to take account of any variation in risk across outcomes. The EAG considered the assessment conducted by the company was appropriate only for the RCT component of the TRuE-V-LTE trial and not the single arm component.

Quality assessment of the trials of ruxolitinib

In general, the EAG agreed with the company's appraisal of the TRuE-V1 and TRuE-V2 trials as assessed using the CRD checklist. In addition to the ratings provided by the company, however, the EAG noted that:

- a minority of people (20%) in the vehicle cream arm showed a meaningful response to treatment as assessed on the F-VASI50 during the double-blind phase of the TRuE-V1 and -V2 trials, even though other active treatments for vitiligo were prohibited. The EAG therefore considered that relative effect estimates from the trial (i.e., the difference between treatment arms) would be more reliable than absolute effects (i.e. the magnitude of the response in the ruxolitinib arm), and that this assumption should apply to all outcomes. Accordingly, the EAG considered that the results from the treatment extension period were at a higher risk of bias.
- The company did not adjust for multiplicity in the analysis, which means that there is an increased risk of a type I error (i.e., incorrectly concluding that there is a statistically significant treatment effect)

- The EAG noted that the upper range in received doses of ruxolitinib and vehicle cream exceeded the recommended dose by a considerable margin. The EAG requested but did not receive detailed information from the company about the dose received by participants in the trial (clarification question B10) and therefore were uncertain how many people who received ruxolitinib in the trials exceeded the dose that is recommended by the product licence. The EAG was uncertain to what extent this would have affected treatment effect estimates in the trials, however considered that this was a potential source of uncertainty.

- The company appraisal did not note that the treatment extension (24 – 52 weeks) was open-label. Open-label trials increase the risk of detection bias, as knowledge of the intervention received can affect the measurement of outcomes.

The company assessment of the TRuE-V1 and TRuE-V2 trials rated the risk of selection bias to be low, on the basis that all outcomes were reported in the trial CSRs. However, the EAG considered the risk of selection bias to be high in the CS, as not all scoped outcomes were presented fully in the main submission (Document B).

For the double-blind phase of the TRuE-V-LTE trial (Cohort A), the EAG did not fully agree with the company's appraisal, for the following reasons:

- The assessment did not take into account in the risks of re-randomising a sub-population of participants selected from the previous trials on the basis of treatment outcome. Participants entering the TRuE-V-LTE were those with no safety concerns after receiving ruxolitinib, as judged by the investigator, and those who had completed the study (which may be influenced by treatment efficacy). This naturally leads to a risk of selection bias in the trial.

- The company's assessment of dropouts from the trial considered differential drop out between arms only and did not consider the high absolute rate of attrition in the trial (approximately 30% in the ruxolitinib arm and 15.5% in the vehicle cream arm). The most common reasons for drop out from the trial were reasons related to treatment outcome. The EAG considered the high rate of missing participants to also represent a risk of bias, in addition to the differential rate between arms.

- The company referred to a 'Table 8' for the methods used to account for missing data. The EAG assumed that this was a typo and the company meant to cite Table 12, which describes the methods for statistical analysis used in the TRuE-V-LTE trial. The methods

described in this table would not be sufficient to account for the missing data in the trial, and are more simplistic than those used in the TRuE-V1 and TRuE-V2 trials (which had minimal missing data). The EAG considered that the outcomes from the TRuE-V-LTE trial were at a high risk of bias because of missing data.

As with the earlier trials, the company rated the TRuE-V-LTE trial as being at a low risk of selection bias as all outcomes were reported in the trial CSRs. However, the EAG considered the trial reporting in Document B of the CS to be at a high risk of selection bias, as results were not fully reported.

For the single-arm cohort of the TRuE-V-LTE trial (Cohort B), as noted in Section 3.1, the company used an inappropriate tool (the CRD checklist for RCTs). Within the timeframe of the EAG appraisal, it was not possible for the EAG to conduct a formal quality appraisal using an appropriate tool. However, informally, the EAG noted the following issues:

- Single-arm trials are subject to a high risk of bias as they cannot control for the possibility that factors other than the treatment may influence treatment outcomes; for example, natural changes in the condition over time.
- There is a risk of selection bias, as participants were selected from the previous trials on the basis of treatment outcome.
- As noted by the company, the trial was open-label, which introduces additional bias, such as detection bias.
- There was a high rate of missing data (>20% participant attrition), which was not accounted for in analyses. The most common reasons for discontinuation from the trials were related to treatment outcome.

Quality assessment of outcomes in the prior therapy subgroup using the TRuE-V1/TRuE-V2

The EAG were aware that the effectiveness data used in the economic model primarily came from the “prior therapy” subgroup. Given that this was the clinical data primarily used in the economic model, the EAG undertook quality assessment for this subgroup using the tool taken from CRD’s guidance for undertaking reviews in health care (Table 10, below). The appraisal

was based on information available to the EAG, which did not include full participant characteristics and outcome data for the previously treated subgroup (Key Issue 2). The EAG concluded that the outcomes linked to the TRuE-V1/TRuE-V2 prior therapy subgroup are at high risk of bias. This was primarily because it was unclear if the treatment arms were similar at the outset, there were no details of how many participants in the subgroup withdrew from the trial or whether this was similar between treatment arms, poor reporting of the outcomes, and a lack of clarity about the analysis used.

Table 10. Quality assessment of the outcomes linked to the TRuE-V1/TRuE-V2 prior therapy subgroup

	TRuE-V1 and TRuE-V2 prior therapy subgroup
Was the method used to generate random allocations adequate?	Participants were centrally assigned to study treatment using an interactive response technology system. Participants were not stratified by prior therapy and this analysis breaks randomisation.
Was the allocation adequately concealed?	Yes, allocation generated by automated system
Were the groups similar at the outset of the study in terms of prognostic factors?	It was unclear if the groups were similar at outset. The company provide the total number of participants in the prior therapy subgroup in Table 32 (Doc B) but no baseline characteristics by treatment arm.
Were the care providers, participants, and outcome assessors blind to treatment allocation?	Yes, double-blind design
Were there any unexpected imbalances in dropouts between groups?	It was unclear how many participants in the subgroup withdrew from the trial or whether this was similar between treatment arms.
Is there any evidence to suggest that the authors measured more outcomes than they reported?	As noted in Key Issue 2, at clarification (QA2) the company were invited to provide evidence for the prior treated subgroup, however this was submitted in a series of inadequately labelled appendix tables including superfluous data, rather than as a transparent submission of selected and pooled estimates from the trials.
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	This was the analysis of a subgroup rather than an ITT analysis. It was unclear what methods were used to account for missing data in the previously treated subgroup, as the data were not provided. However, multiple imputation was used for the ITT analysis.

Abbreviations: ITT, intention to treat

3.2.3. Description and critique of the results of the studies

3.2.3.1. Clinical effectiveness results

Clinical effectiveness data in the CS was largely based on pooled data from the TRuE-V1[25] and TRuE-V2[19] trials. Overall, the EAG considered that the company's evidence submission was poor and lacking in transparency. Not all scoped outcomes were presented by the company. In some cases, the company referred the EAG to documents produced by the EMA in the SmPC report of ruxolitinib(EMA)[26] (provided in Appendix C of the CS), though some of these data were only available in poor resolution figures and lacked detail (note that SmPC reports are not produced with the aim of being submitted for appraisal within the HTA process). Some data required by the NICE decision problem were only available in trial CSR documents and appendices. Notably, this included clinical effectiveness outcome data for the previously treated subgroup, which were not in the CSR documents provided by the company (though the EAG requested that all CSR files, including tables and appendices be submitted [clarification question C2]) but were submitted by the company in a series of files at clarification that appeared as if they were originally an appendix to the CSRs. These were poorly labelled, which led to uncertainty about the data source. In some cases, the EAG attempted to calculate data for the pooled trial population from the individual CSRs, to aid comparability with other trial outcomes, but noted that these data would not consider missing values analysis and may be based on different analysis populations than data reported in the CS.

Overall, the poor reporting standard of clinical effectiveness evidence in the CS and in subsequent submissions from the company undermined the ability of the EAG to fully appraise the clinical effectiveness of ruxolitinib. The EAG also considered there to be a risk of selection bias in the CS (Section 3.2.2.6). In particular, the EAG was concerned about the reliability of data for the previously treated subgroup, which is the company's chosen indication for ruxolitinib and is the population used in its economic evaluation (Key Issue 2).

In this section, the EAG provides a summary of its appraisal of the clinical effectiveness evidence for all scoped outcomes in the main trial population (i.e., regardless of previous treatment status). Due to the reasons outlined above, the appraisal may have gaps or be uncertain in places.

Change in facial vitiligo

The response rate for the ITT population from the pooled trials on the F-VASI at 50%, 75% and 90% is shown in Table 11 alongside the mean change in the F-BSA scale. The EAG noted that a significant minority (20%) of people in the vehicle cream arm showed a >50% reduction in facial vitiligo (F-VASI) in the 24 weeks from baseline, even though active treatments for vitiligo were prohibited during the trial. The EAG was uncertain what would cause this effect, as it was unclear to what extent the F-VASI measure would be susceptible to subjectivity bias. Given the high rate of response in the vehicle cream arm, the EAG considered that the relative treatment effects for all outcomes during the double-blind trial phases would be most reliable for determining the effectiveness of ruxolitinib (i.e., as opposed to the absolute data in each arm). The EAG also considered that outcomes from the uncontrolled trial phases (the treatment extension and Cohort B analysis of the LTE) may best be interpreted with caution.

The data demonstrated that ruxolitinib was more effective than vehicle cream at reducing facial vitiligo, as assessed across all outcomes and accounting for imprecision in the treatment effects. In the 24 weeks from baseline, 21% more people receiving ruxolitinib achieved a response that was above the threshold considered by the EAG's clinical expert to be clinically meaningful to people with vitiligo (i.e. F-VASI 75%) compared to vehicle cream. The effect of ruxolitinib on the F-VASI increased further between week 24 and 52, with more people initially allocated to ruxolitinib achieving a greater level of response by the 52-week timepoint. By 52 weeks, half of all those initially allocated to ruxolitinib had achieved a response above the threshold considered to be clinically meaningful (i.e. F-VASI 75%). This was 40% more people than had achieved a response with vehicle cream at 24 weeks.

Figures provided by the company (Figure 10 and 11, CS Doc B p.66 & 70) showed an increasing response rate to treatment over time; for example, the number of people achieving F-VASI75 following ruxolitinib increased from 5.3% at 8 weeks (no difference with vehicle cream) to 31.0% at 24 weeks. In the treatment extension, response rates in those switching to ruxolitinib appeared to follow the same trajectory as those allocated to ruxolitinib in the initial double-blind phase.

Mean change in F-VASI score in the two trials was shown in figures only in the SmPC and in the respective trial CSRs (showing change in F-VASI scores with respective standard errors). The EAG was unable to identify specific data, including mean, median, min and max change in any of the documents supplied by the company (the trial CSRs received by the EAG reported the

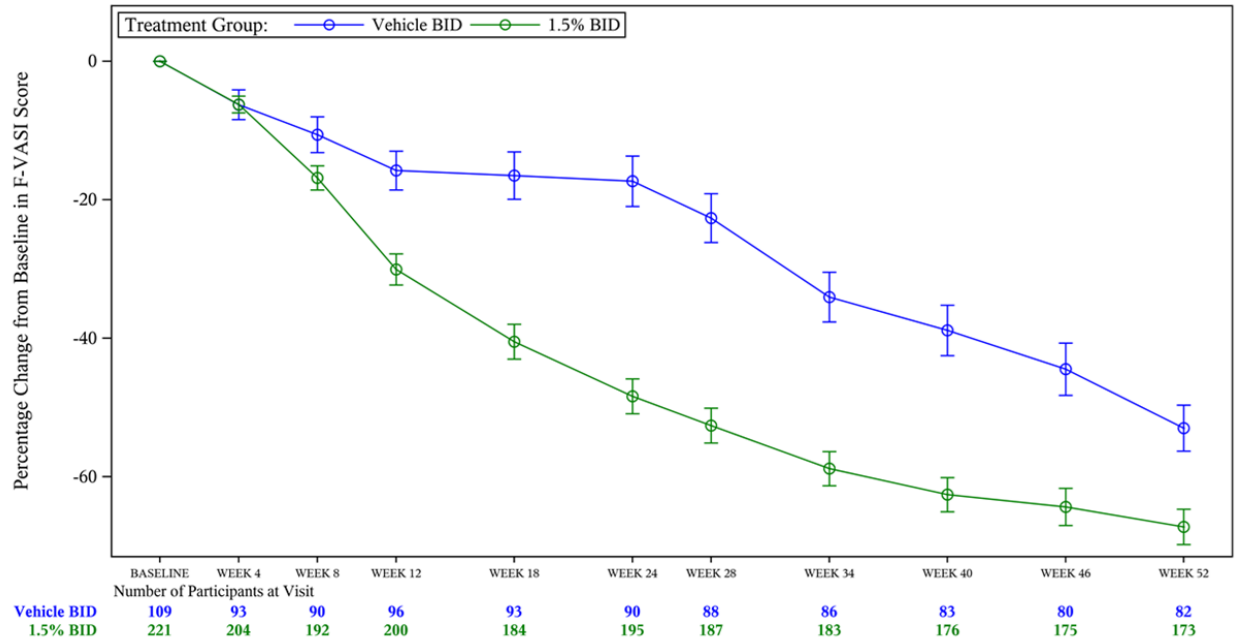
data in figure form only). The figures are reproduced below (Figure 3). The figures showed a steady increase in the treatment effect of ruxolitinib over time, though standard error bars suggested that this effect varied meaningfully across the population. This was consistent with the response rate data, showing that many participants in the trials did not experience a clinically meaningful response to ruxolitinib. The EAG also noted that the curve gradient began to plateau from week 34 onwards, suggesting that limited further improvements in facial vitiligo may occur beyond this timepoint. This effect was also visible in mean T-VASI scores (see next section and shown on p.166 of the SmPC report)[1].

Table 11: Change in facial vitiligo outcomes based on pooled data from TRuE-V1 and TRuE-V2

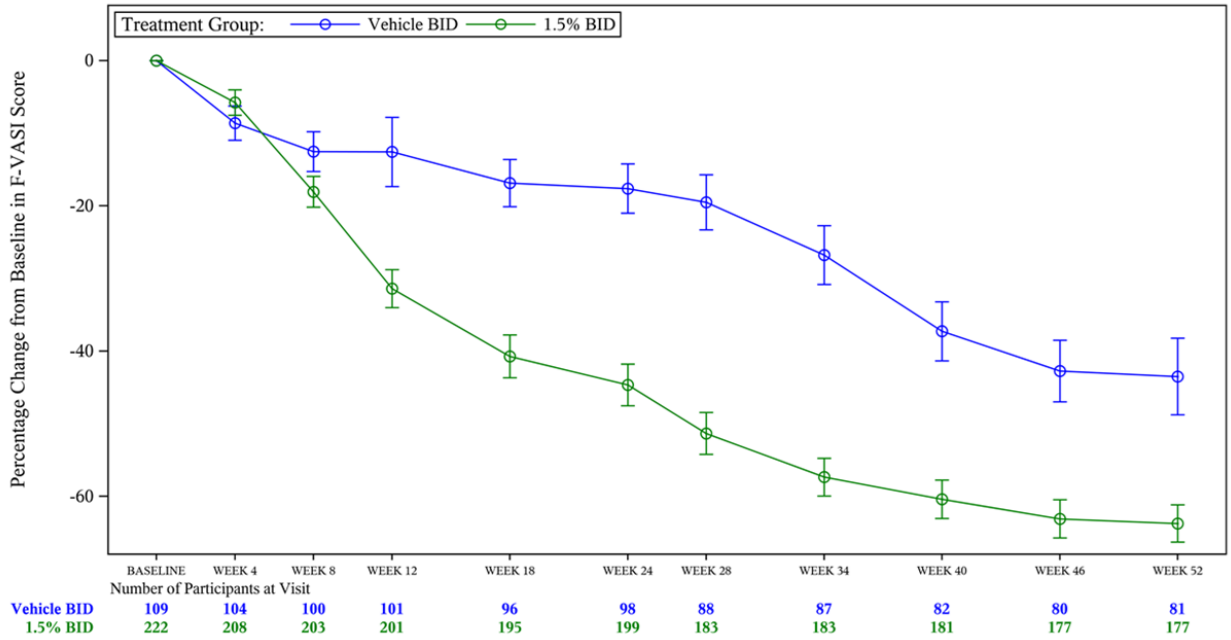
	Response rate						% LSM change	
	F-VASI50		F-VASI75		F-VASI90		F-BSA	
	Vehicle (N=218)	Rux (N=443)	Vehicle (N=218)	Rux (N=443)	Vehicle (N=218)	Rux (N=443)	Vehicle (N=218)	Rux (N=443)
Week 24	19.6% (SE 2.89)	51.7% (SE2.46)	9.6% (SE2.17)	30.7% (SE2.29)	1.9% (SE1.01)	16.0% (SE1.83)	-7.9% (95%CI -13.02, -2.69)	-27.8 % (95%CI -31.29, -24.41_)
Difference between arms		32.2% (95%CI 24.6, 39.7)		21.1% (95%CI 14.9, 27.3)		14.2% (95%CI 10.1, 18.3)		-20.0% (95%CI -26.2, -13.8)
OR (95%CI)		4.40 (2.92, 6.65)		4.17 (2.43, 7.14)		10.33 (3.31, 32.2)		NA
	Vehicle – Rux (N=163)	Rux – Rux (N=350)	Vehicle – Rux (N=163)	Rux – Rux (N=350)	Vehicle – Rux (N=163)	Rux – Rux (N=350)	Vehicle – Rux (N=163)	Rux – Rux (N=350)
Week 52	52.8%	74.6%	28.2%	50.3%	14.1%	30.3%	-26% (-22, -30)	-42.5% (-41, -44)

Note: Response rate data is estimated based on the company's analyses (described in section B.2.4 of the CS). All data shown are using the company's multiple imputation approach. Data at 52 weeks was taken from figures provided in the EMA SmPC report. Data for F-BSA was approximate based on figure curves and error bars and therefore may be inaccurate. Data from study site 710 were removed from all data points.

Figure 3: Change (mean ±SE) in F-VASI in TRuE-V1 and TRuE-V2



Note: Data from study site 710 were excluded.



Note: Data from study site 710 were excluded.

Note: source EMA SmPC, appendix C of the CS (p.115)

Physician- and patient-reported assessments of improvement in vitiligo were not reported in the CS. The company cited the SmPC report[1], which presented the results separately for the two trials. The EAG calculated a naïve pooling of the trials, shown in Table 12. The results were consistent with results from the F-VASI scales.

Table 12: Physician- and Patient-reported improvement in facial vitiligo from TRuE-V1 and TRuE-V2

	Vehicle cream (N=109)	Ruxolitinib (N=221)
F-PhGVA score of clear (0) or almost clear (1)		
Week 24	9.04%	30.75%
Week 40	21.74%	40.28%
Week 52	27.61%	42.82%
F-PaGIC V score of very much improved (1) or much improved (2)		
Week 24	7.98%	42.64%
Week 40	32.73%	50.28%
Week 52	38.04%	53.14%

Source: calculated based on data reported in the SmPC[1]

Note: data from study site 710 were excluded

Maintenance of response in facial vitiligo

The EMA SmPC[1] report provided shift summary data for those receiving ruxolitinib across both trials for week 24 to 52, which showed how participants' treatment response changed between these time points. These data are reproduced below (some participants were missing from these data [n=44], implying that multiple imputation was not used). The reasons for missingness were not reported and so it's possible that some data were missing due to treatment outcome (see critical appraisal of the included trials, section 3.2.2.6). The data showed the following:

- A minority of people (10.3%) experienced a deterioration in treatment response between 24- and 52-weeks.
- Approximately a third of people (38.8%) remained in the same response category between week 24 and week 52

- In all categories under F-VASI-90 (the highest response), more people (51.5%) showed a further improvement in response between week-24 and -52 than remained in the same category.

Table 13: Shift summary of maintenance response on F-VASI (ITT pooled population) from week 24 to week 52 (ruxolitinib arm)

Response at Week 24		Response at Week 52					
Value	n (%)	< F-VASI25	F-VASI25 to < F-VASI50	F-VASI50 to < F-VASI75	F-VASI75 to < F-VASI90	F-VASI90	Missing
Pooled analysis							
< F-VASI25	131 (33.2)	47 (35.9)	21 (16.0)	30 (22.9)	7 (5.3)	8 (6.1)	18 (13.7)
F-VASI25 to < F-VASI50	59 (15.0)	0	9 (15.3)	21 (35.6)	12 (20.3)	4 (6.8)	13 (22.0)
F-VASI50 to < F-VASI75	82 (20.8)	1 (1.2)	8 (9.8)	26 (31.7)	22 (26.8)	19 (23.2)	6 (7.3)
F-VASI75 to < F-VASI90	58 (14.7)	0	2 (3.4)	7 (12.1)	22 (37.9)	26 (44.8)	1 (1.7)
F-VASI90	64 (16.2)	1 (1.6)	0	1 (1.6)	7 (10.9)	49 (76.6)	6 (9.4)

Note 1: Data from participants enrolled at Site 710 in Study INCB 18424-307 were excluded.
Note 2: The analysis was conducted in the ITT population for participants in the ruxolitinib 1.5% cream BID group with nonmissing F-VASI scores at Week 24.

Source: EMA SmPC report, p. 124

The company reported shift summary data for the TruE-LTE trial between week 52 to week 104 for those receiving ruxolitinib who did not respond to treatment (i.e., those who received ruxolitinib in the original trials who did not respond, entered cohort B and continued to receive ruxolitinib; CS Doc B p.121). These data showed that continuing improvement in F-VASI was possible beyond week 52, but that a deterioration in response was also possible. Again, some data were missing from these data ([redacted]%), and as reasons for missingness could include reasons related to treatment efficacy, the precise rates of movement between response category were uncertain.

Facial vitiligo outcomes for responders (Cohort A) at week 104 in the TruE-V LTE trial are shown in Table 14. The EAG had concerns about these data on the basis that a reasonable minority of participants in both arms were censored due to treatment discontinuation: 23.2% in the vehicle cream arm and 12.7% in the ruxolitinib arm. The EAG was unable to identify the reasons for discontinuation of these participants from the information provided by the company,

however in general the biggest reason for discontinuing from the trials was loss to follow up and withdrawal by participant. The EAG considered it plausible that these discontinuations would not be random, but participants would have discontinued due to the efficacy or safety of treatment. The EAG considered that the number of participants missing from this analysis at 104 weeks was sufficient to potentially bias the results, and therefore considered that the data should be interpreted with caution. To account for this, the EAG calculated relapse rates in each arm to include those who discontinued the trial (i.e. assuming all who discontinued the trial relapsed) and/or those who received rescue medication (i.e. those who received ruxolitinib to maintain a response; note that the EAG was unclear how this was administered in those who were continuing with ruxolitinib during the LTE). These data are also shown in Table 14.

Based on the number of people shown to have experienced a relapse and according to the company's calculation (<FVASI75), 14.5% of people who responded to ruxolitinib relapse within 2 years while still receiving treatment. However, when including all those who discontinued treatment, 27.3% of people who continue treatment with ruxolitinib will relapse within 2-years, and this was 30.9% when also including those who received rescue medication. Continuing with ruxolitinib after achieving a response was nevertheless associated with a reduced risk of relapse: twice as many people who discontinued ruxolitinib experienced a relapse than those who continued with treatment. The rate of relapse after discontinuing treatment was 28.6%, or 60.7% if including all those who discontinued the trial and received rescue treatment.

At the time of submission, the median time to relapse was not estimable in either group and the relative hazard for relapse was highly imprecise. However, the EAG was persuaded that continuing with ruxolitinib was likely to reduce the risk of relapse compared to discontinuation.

Table 14: F-VASI75 at week 104 for those who responded in the TRuE-V1 and TRuE-V2 trials

	Responders	
	Switched to vehicle (N=56)	Continued with rux (N=55)
<F-VASI75 (relapse); N (%)	16 (28.6%)	8 (14.5%)
Time to F-VASI75 (relapse); days (95%CI)	NE (238.0, NE)	NE (NE, NE)
HR (95%CI)		0.422 (0.18, 0.99)
<F-VASI75 (relapse) including those censored for tx discont	29 (51.8%)	15 (27.3%)
<F-VASI75 (relapse) including those censored for tx discont	34 (60.7%)	17 (30.9%)

	Responders	
and those who received rescue therapy		

Abbreviations: discont, discontinuation; F-VASI, Facial Vitiligo Area Scoring Index; HR, hazard ratio; NE, not estimable; rux, ruxolitinib

Change in total vitiligo

Only the rate of people meeting T-VASI50 was reported in the CS. To determine the response rates at other thresholds, the EAG identified data reported separately from the trial CSRs. The EAG has attempted to calculate response rates where feasible, but these do not account for censoring, and no continuous or variance data were available. The data available are shown in Table 15.

As with facial vitiligo outcomes, a minority of people in the vehicle cream arm reported meaningful improvements in bodily vitiligo, though this was lower than for F-VASI – approximately 6% of people who received vehicle cream were reported to have experienced a >50% reduction in total vitiligo during the 24-week DB period.

Response rates were lower for total vitiligo than facial vitiligo: 6.1% and 36.4% of people receiving ruxolitinib achieved a meaningful response in total vitiligo (T-VASI75) at 24- and 52-weeks, respectively, compared to 30.7% and 50.3% in F-VASI. The EAG considered it plausible either that (a) there may be different mechanisms involved in bodily and facial vitiligo, and therefore outcomes may not be well correlated, and/or (b) that bodily vitiligo was slower to change and that further improvements in bodily vitiligo may be seen with longer follow-up. The EAG was aware of evidence that some parts of the body (e.g., hands, feet, lips) may be less likely to respond to treatment for vitiligo than the face and trunk, but was unsure how established this effect is, whether it would be consistent across treatment types, and whether the effect would be sufficient to explain the difference between F-VASI and T-VASI outcomes. As shown in the next section, a similar proportion of people showed an improvement in response between 24-weeks to 52-weeks on the T-VASI as F-VASI, and on the whole the EAG did not consider there to be evidence of a delayed treatment effect for bodily vitiligo.

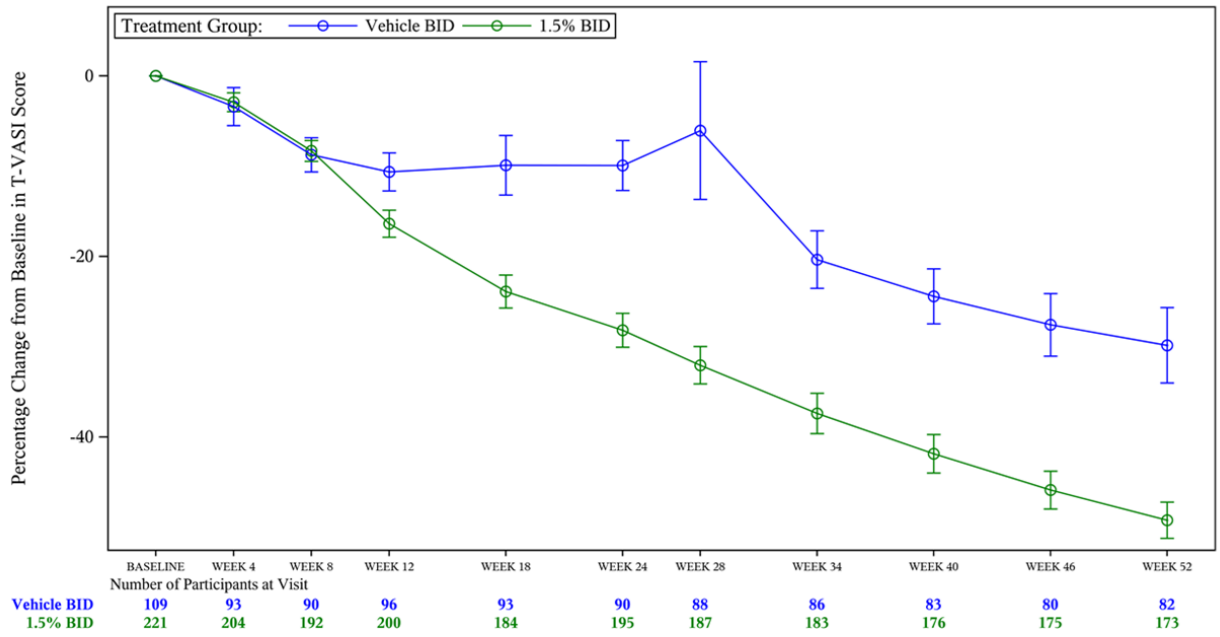
Table 15: Change in bodily vitiligo outcomes based on pooled data from TRuE-V1 and TRuE-V2

	Response rate					
	T-VASI50		T-VASI75		T-VASI90	
	Vehicle (N=218)	Rux (N=443)	Vehicle (N=218)	Rux (N=443)	Vehicle (N=218)	Rux (N=443)
Week 24	5.8% (SE 1.64)	21.9% (SE 2.04)	1.8%	6.1%	0%	0.68%
Difference between arms		16.1 (95%CI (10.9, 21.2)		NR		NR
OR (95%CI)		4.55 (2.42, 8.58)		NR		NR
	Vehicle – Rux (N=163)	Rux – Rux (N=350)	Vehicle – Rux (N=218)	Rux – Rux (N=443)	Vehicle – Rux (N=218)	Rux – Rux (N=443)
Week 52	27.0%	51.1%	7.3%	36.4%	1.8%	4.5%

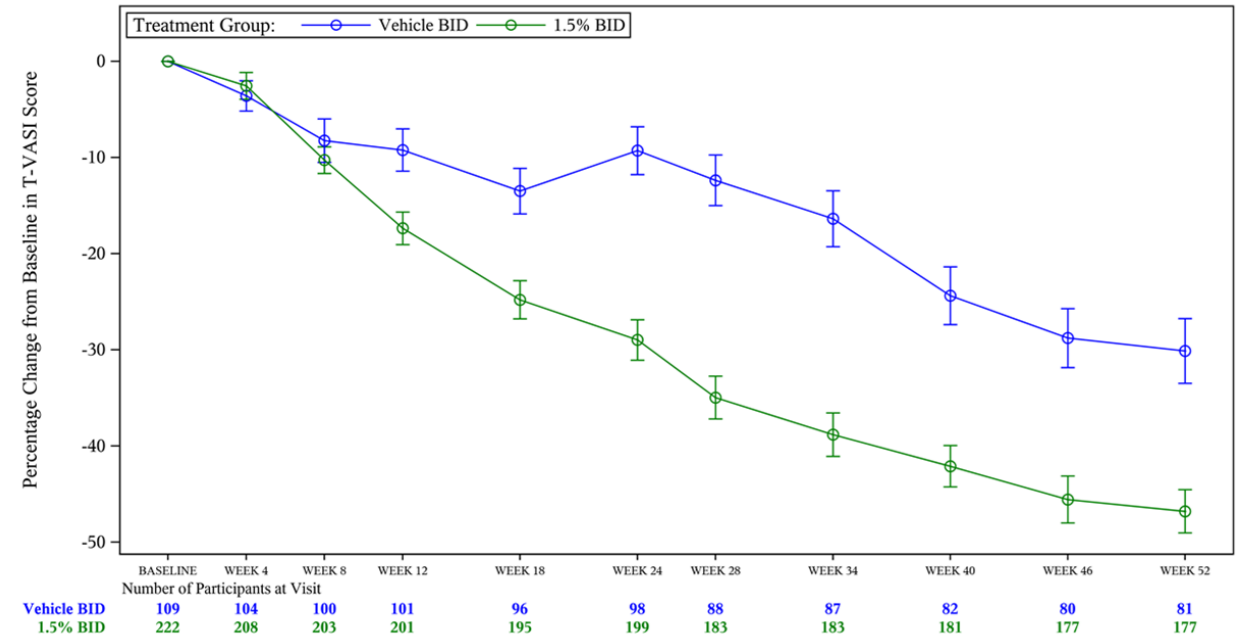
Note: week 52 data for T-VASI75 and T-VASI90 was taken from the CSR and was only available as a % from the ITT population using the company's multiple imputation analysis.

Mean change in T-VASI scores over the two trials was reported in the SmPC report and is shown in Figure 4.

Figure 4: Change (mean ±SE) in T-VASI in TRuE-V1 and TRuE-V2



Note: Data from study site 710 were excluded.



Note: Data from study site 710 were excluded.

Maintenance of response in total vitiligo

The shift summary data between week 24 and week 52 for T-VASI were reported in the EMA SmPC report. Of those who showed a response of ≥T-VASI-25 at 24 weeks, only 6.7% showed signs of relapse (reduction in response category) by 52-weeks. A third of participants (35.6%) remained in the same response category, and 50.3% of participants with a response <T-VASI90 at 24 weeks showed an improvement in response by 52-weeks.

Table 16: Shift summary of maintenance response on T-VASI (ITT pooled population) from week 24 to week 52 (ruxolitinib arm)

Response at Week 24		Response at Week 52					
Value	n (%)	< T-VASI25	T-VASI25 to < T-VASI50	T-VASI50 to < T-VASI75	T-VASI75 to < T-VASI90	T-VASI90	Missing
< T-VASI25	198 (50.3)	75 (37.9)	58 (29.3)	33 (16.7)	6 (3.0)	1 (0.5)	25 (12.6)
T-VASI25 to < T-VASI50	104 (26.4)	4 (3.8)	29 (27.9)	46 (44.2)	11 (10.6)	1 (1.0)	13 (12.5)
T-VASI50 to < T-VASI75	66 (16.8)	0	3 (4.5)	24 (36.4)	28 (42.4)	5 (7.6)	6 (9.1)
T-VASI75 to < T-VASI90	23 (5.8)	1 (4.3)	1 (4.3)	4 (17.4)	9 (39.1)	8 (34.8)	0
T-VASI90	3 (0.8)	0	0	0	0	3 (100.0)	0

Note: Data from participants enrolled at Site 710 in Study INCB 18424-307 were excluded.

Cosmetic acceptability

Response to treatment as measured by the VNS (a score of 4 or 5 indicating their vitiligo is no longer noticeable or a lot less noticeable) was reported in the CS. The data showed that a third of people who received ruxolitinib in the TRuE-V1 and TRuE-V2 trials considered their vitiligo to have become a lot less noticeable after 52 weeks' of treatment.

	VNS score 4 or 5	
	Vehicle (N=218)	Rux (N=443)
Week 24	4.2% (SE1.45)	22.5% (SE2.09)
Difference between arms		18.3% (95%CI 13.3, 23.2)

	VNS score 4 or 5	
OR (95%CI)		6.52 (3.11, 13.67)
	Vehicle – Rux (N=163)	Rux – Rux (N=350)
Week 52	16.6%	36.3%

Abbreviations: OR, odds ratio; rux, ruxolitinib; SE, standard error; VNS, vitiligo noticeability scale

Health-related quality of life

The CS, Document B, did not present HRQoL data evaluated in the trials. The company referred to the EMA SmPC report[1] (CS, appendix C), although this report did not contain data for the DLQI and the CDLQI, just stated that no changes in either outcome were observed over time (EMA SmPC, p.118). The trial CSRs[9,12,18,19] reported DLQI and CDLQI data at baseline and follow-up. Amongst adults, the majority of people reported that their vitiligo had no effect or a small effect on their lives at baseline as assessed using the DLQI. There was no change in DLQI or CDLQI over the trials.

Change in VitiQoL scores (a vitiligo-specific HRQoL measure) was reported separately for each trial in the SmPC report, and no statistically significant difference in scores was reported between groups at the end of the double-blind phase (week 24). The absolute change in HRQoL increased between week 24 and week 52[9,12,18,19], but no statistical tests were conducted to determine if the change from baseline was statistically significant. The company did not report a validated clinically minimally important difference for this measure, and the EAG was unable to identify one during its appraisal. As a consequence, the EAG was unable to determine if participants in either arm showed a clinically meaningful change in VitiQoL during the trials). However, using an arbitrary threshold of 10%, improvements in VitiQoL were <10% at week 24 and marginally above 10% at 52 weeks for those initially assigned to ruxolitinib in TRuE-V1[7] and <10% at all timepoints in TRUE-V2[8](though the EAG highlight that these rates do not account for change in the vehicle cream arm). Variance around VitiQoL scores was extremely wide, however: mean change from baseline at 52 weeks in those originally assigned to ruxolitinib was [redacted] in TRUE-V1, suggesting that effects of ruxolitinib on VitiQoL were extremely varied across the population. The company did not report change in HRQoL scores for those who reported that their condition had a meaningful impact on their lives at baseline.

Psychological wellbeing

At baseline, means scores on the HADS anxiety and depression subscales were within normal range[27] (i.e. not indicative of clinical anxiety or depression; reported in the trial CSRs[9,12,18,19]). The company stated that there was a "numerically greater improvement" in the HADS total score of depression and anxiety (Doc B, p.75). However, not only was this finding not statistically significant, but the 'numerical change' was well under published thresholds for a clinically meaningful change in HADS in any population[28]. The EAG therefore agreed with the assessment of the EMA that there was no difference in HADS score between those receiving ruxolitinib and vehicle cream at 24 weeks. There was also no benefit of ruxolitinib on HADS at 52-week follow-up. The company did not report change in HADS for those who reported clinically significant symptoms of anxiety and depression at baseline.

The EMA SmPC report[1] also reported no meaningful difference in outcomes on the WHO-5 (a measure of general wellbeing) between trial arms.

Subgroup analyses

Subgroup analysis of F-VASI75 presented by the company (CS Doc B, p.82) showed a differential treatment effect according to participant age (larger effect in adolescents than adults) and facial BSA at baseline (larger effect in those with greater facial vitiligo). Clinical advice to the EAG was that these findings would be expected, given that these groups tend to show better outcomes following all treatments for vitiligo. There was no difference in treatment effect between participants with Fitzpatrick scale Type 1/2 and Type 3/4/5/6.

Data for some outcomes was provided for trial participants who had previous received treatment at clarification (question C2). These data appeared to be excerpts from the appendices of the trial CSRs, though these tables were not provided to the EAG in the appendices of CSRs provided earlier in the appraisal (though these and all data tables were requested by the EAG). Within the timeframe of the appraisal, it was not possible for the EAG to review all these documents, however the EAG considered the documents that related to the primary outcome of the trials (F-VASI75). The three files for this outcome were not adequately labelled and the sample sizes reported in the tables did not clearly match the trial data to confirm identification, however the EAG assumed the following:

- File 'T_1_1_1_1_FVASI75.RTF' reports data for one of the TRuE-V1 and TRuE-V2

- File 'T_1_1_1_2_FVASI75.RTF' reports data according to whether participants received any previous treatment for the pooled TRuE-V1 and TRuE-V2 trials
- File 'T_1_1_1_3_FVASI75.RTF' reports data according to whether participants received previous TCI or TCS

The data assumed to be based on the pooled trials showed that those who had previously received treatment showed a very slight increased chance of a response to ruxolitinib compared to the broader population ([redacted] vs 30.7% at 24 weeks; [redacted] vs 51.6% at 52 weeks). The company did not report a formal subgroup analysis to compare response between those who did and did not receive previous treatment, however the EAG assumed that there would be no statistically significant difference between groups. As described in Key Issue 2, there is significant uncertainty over the data used by the company to represent the previously treated subgroup.

Adverse effects

Ruxolitinib was associated with a small increase in the risk of adverse events compared to vehicle cream. Mostly these were mild adverse events but there was an increased risk of adverse events affecting the treated area, including acne, pruritus, erythema and rash. The EAG considered that these would not contribute to major health concerns or healthcare resource use, though considered that people using ruxolitinib who experience these events may be more likely to discontinue treatment or else change the application of ruxolitinib to another area of the body. The rate of adverse events increased between the 24-week and 52-week timepoint, suggesting that new events may emerge with longer exposure.

There was also a small increase in the rate of serious adverse events in those who received ruxolitinib. The trial investigators determined that none of these events were related to treatment. Event rates were extremely low and with no obvious pattern that was suggestive of a particular risk with ruxolitinib.

Oral ruxolitinib has been associated with an increased risk of nonmelanoma skin cancer[29] (NMSC) in other skin conditions. The EAG was unclear how dosing between the topical and oral formulations of ruxolitinib compared, though the company reported no skin cancer events in the TRuE-V1 and V2 trials, or in the TRuE-V-LTE (as reported in the CS and trial CSRs). However, the SmPC[1,26] report notes that ten participants with vitiligo receiving ruxolitinib across the broader evidence base (including trials not reported in the CS) experienced a non-melanoma

skin neoplasm, most commonly (n=3) basal cell carcinoma. The SmPC for ruxolitinib[26] noted that a causal relationship with ruxolitinib has not been identified, though "4/5 patients had NMSC at an application site" (p.168) and they considered that the follow-up of the ruxolitinib trials in vitiligo was insufficient to determine whether NMSC may develop over time. The EMA, MHRA and clinical advice to the EAG concurred that people who receive treatment with ruxolitinib should be monitored for skin cancer, pending further evidence.

3.3. Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

The company conducted a SLR to identify relevant clinical trial evidence for the submission. The results of SLR were used to assess the feasibility of a robust indirect treatment comparison (ITC) to estimate the relative efficacy of ruxolitinib versus other therapies. A summary of this process is reported in Section B.2.9 of the CS, with more detail of the methods presented in Appendix D1.2. The population, intervention, comparator, outcomes, and study design (PICOS) criteria used in the ITC are presented in Table 12, Appendix D.

A total of 253 studies were included in the SLR and were screened for inclusion in the feasibility assessment (FA). Twenty-four potential comparator studies and four studies related to ruxolitinib cream were included in the ITC FA. The screening process was reported in Table 13 (Appendix D) but specific reasoning for excluding studies was not presented. After the feasibility assessment, the company concluded that there was an insufficient evidence base to robustly compare the efficacy of ruxolitinib to existing therapies. They reasoned that the lack of comparable studies was partly due to an evolving set of tools that were used to evaluate vitiligo. In addition, they noted that most of the clinical studies were of low methodological quality.

3.4. Additional work on clinical effectiveness undertaken by the EAG

The EAG conducted an assessment of studies included in the company's SLR that could plausibly be included in an ITC. The majority of the studies identified by the company were small, often fewer than 50 participants, and the EAG agreed that there was between-study heterogeneity in terms of study design and patient population characteristics.

However, the HI-Light Vitiligo Trial[30] was a large, placebo controlled RCT, conducted in the UK. Participants were randomised to either dummy NV-UVB plus TCS (TCS group), NV-UVB plus vehicle cream (NB-UVB group), or NV-UVB plus TCS (combination group). The NV-UVB used was a home-based handheld narrowband ultraviolet B (NB-UVB). As noted in Section 2.4, the

EAG considered that the relevant comparator for this submission was existing 2nd line treatment options, including NV-UVB with or without TCS, in line with the treatment pathway published by BAD[21]. The EAG considered that the HI-LIGHT trial was a highly relevant evaluation of treatments for NS vitiligo, and noted that the company used published evidence from the HI-Light Vitiligo Trial[30] within a multistep process to estimate patient EQ-5D-3L utility values to assign to health states in the company's economic model (see Section 4.2.8).

The EAG independently considered the feasibility of conducting an ITC to compare ruxolitinib to NV-UVB plus TCS utilising the HI-Light Vitiligo Trial. After an appraisal of the available evidence base, the EAG considered that a network meta-analysis (NMA) could potentially be conducted using Eleftheriadou 2014[31], the pilot Hi-Light trial, to connect TRuE-V1[18] and TRuE-V2[19] to the HI-Light Vitiligo Trial[3] in the analysis. The EAG also considered that the company may have been able to utilise individual patient data (IPD) from the pooled TRuE-V1 and TRuE-V2 trials to perform an unanchored matching-adjusted indirect comparison to the relevant arm in the HI-Light Vitiligo Trial. The EAG appraisal of the feasibility of these options is described in the following sections. Details of the studies considered by the EAG in its assessment, including inclusion and exclusion criteria, baseline demographic and disease characteristics, and potential outcomes are presented in Table 34 in Appendix A.

3.4.1. Network meta-analysis

The EAG assessed whether conducting an NMA utilising Eleftheriadou 2014, the pilot Hi-Light trial, to connect TRuE-V1and TRuE-V2 to the HI-Light Vitiligo Trial could produce a credible estimate of effect.

A key limitation to this approach would be the small size of Eleftheriadou 2014, the pilot Hi-Light trial. This pilot trial included 19 participants in the NB-UVB arm and 10 participants in the placebo arm. This led to treatment groups that were not well matched in terms of their baseline or disease characteristics. The EAG was also concerned that there were systematic differences in the participants recruited to the comparator trials and the ruxolitinib trials. The age, and BSA involvement inclusion criteria varied between TRuE-V1/TRuE-V2 and the comparator studies. Also, the participants recruited to the HI-Light Vitiligo Trial had more progressive disease compared to the TRuE-V1and TRuE-V2 trials, where only 175 (26%) were reported to have progressive disease. Given the limitations noted, the EAG did not consider an NMA to be a robust approach to estimate the effectiveness of ruxolitinib in comparison to NV-UVB with or without TCS.

3.4.2. Unanchored matching-adjusted indirect comparison

The EAG assessed whether performing an unanchored matching-adjusted indirect comparison (MAIC) to the combination arm in the HI-Light Vitiligo Trial could produce a credible estimate of effect.

An unanchored MAIC approach does not require the use of the small Eleftheriadou 2014 in the analysis. It would allow for the analysis to be adjusted to match variation in baseline characteristics reported in the HI-Light Vitiligo Trial, and for which TRuE-V1/TRuE-V2 has IPD. The EAG noted that there were baseline characteristics not reported in the HI-Light Vitiligo Trial that cannot be matched utilising this analysis. This included disease status and mean T-BSA (% of the total body involved). This is in addition to the limitation inherent to all MAICs that despite the use of IPD to reduce observed cross-trial differences, unobserved differences may result in residual confounding.

The principal limitation of using unanchored MAIC for this analysis was variations in the outcomes reported in the trials. The primary outcome reported in the TRuE-V1/TRuE-V2 trials was F-VASI75, which constitutes an improvement of at least 75% from baseline in the F-VASI. The company also reported T-VASI75, which constitutes an improvement of at least 75% from baseline in the T-VASI. The repigmentation outcome reported in the HI-Light Vitiligo Trial is > 75% repigmentation using digital images taken at baseline and at 9 months of a single “target patch” per person. This target patch has “active” vitiligo and is therefore new or has changed over the past 12 months.

The EAG’s clinical expert noted that VASI measures depigmentation on the whole body (T-VASI) or the whole face (F-VASI). Thus, it is a global measure of repigmentation rather than targeting a single active patch and it could include patches that are stable and patches that are progressive. The EAG noted that the target patch was active and therefore more likely to be classed as progressive. The EAG’s clinical expert confirmed that active patches were thought to be more responsive to treatment.

Given substantial differences between the outcomes reported in the TRuE-V trials and the HI-Light Vitiligo Trial, the EAG did not consider an unanchored MAIC to be a robust approach to estimate the effectiveness of ruxolitinib in comparison to NB-UVB plus TCS.

3.5. Conclusions of the clinical effectiveness section

Overall, the EAG considered that the presentation of clinical effectiveness data in the CS was poor and lacked transparency, which prevented a full appraisal of the clinical effectiveness of ruxolitinib. The EAG was particularly concerned about the omission of clinical effectiveness evidence in the population subgroup that was used in the company's economic model (previously treated) and the omission of several completed trials of ruxolitinib from the CS. From the evidence appraised by the EAG, the EAG considered that:

- A significant minority of people in the clinical trials showed a clinically meaningful response in facial vitiligo with ruxolitinib over and above vehicle cream. This response was above the threshold considered by the EAG's clinical expert to be meaningful for people with vitiligo. Vitiligo patches on the face are particularly distressing for people with vitiligo, and the EAG considered that the improvements in facial vitiligo for these participants would be meaningful to them.

- The treatment effect of ruxolitinib for total vitiligo was lesser than that of facial vitiligo, with fewer participants showing a response in TVASI. The EAG was unable to explain the reduced efficacy for TVASI and considered there to be no evidence that the treatment response for bodily vitiligo would be slower to emerge. While the licence for ruxolitinib was limited to people with vitiligo affecting the face, there was no restriction on where people who receive ruxolitinib can apply the cream (up to 10% of BSA). As vitiligo affecting the face was one of the symptoms that people with vitiligo reported to be most distressing, the EAG considered it reasonable to assume that most people with facial vitiligo would apply the cream to their faces. However, the EAG noted that variation in the location of the cream may affect the clinical benefits experienced.

- The effect of ruxolitinib increases over time, with some variation across people in the speed and magnitude of response experienced. Clinical advice to the EAG was that dermatologists will typically continue treatments for vitiligo when people show >20% pigmentation change every 3-4 months. As more than half of people treated with ruxolitinib did not show a clinically meaningful response as compared to vehicle cream, the EAG considered that a strategy to allow a response to develop while discontinuing those who will not experience a treatment benefit would be optimal for prescribing ruxolitinib.

- Trial participants did not show overall benefits of ruxolitinib for HRQoL or psychological wellbeing. The EAG considered it plausible that people who experienced significant improvements in their vitiligo may experience a meaningful benefit in these outcomes, but these data were not presented by the company.
- Ruxolitinib appeared to be associated with a low risk of adverse events, with the most common adverse events being mild in nature. Nevertheless, the EAG considered that the types of adverse events reported may affect treatment use; for example, people may choose not to apply ruxolitinib to their face if they experience acne, and this in turn may affect treatment efficacy.
- The EAG considered that the effectiveness of ruxolitinib in the subgroup of people who had previously received treatment was uncertain. The EAG had no evidence to believe that treatment effects would be different in those who had previously received treatment, however considered that an appraisal of these data would be useful to reduce this uncertainty.
- The relative effectiveness of ruxolitinib as compared to other treatments for vitiligo was an ongoing source of uncertainty in the appraisal. Following an appraisal of the evidence base identified in the company's SLR, the EAG agreed with the company's conclusion that a statistical comparison of ruxolitinib with relevant 2nd line treatment options using either an NMA or a MAIC was not feasible and/or would not be useful for decision-making. A naïve comparison of clinical outcomes between people who received ruxolitinib in the TRuE-V trials and outcomes reported in a large, UK based trial[30] of NB-UVB therapy and combination TCS and NB-UVB therapy suggested that more people may respond to ruxolitinib than either of the other treatments. However, without a head-to-head comparison, any conclusions about the relative effectiveness of ruxolitinib would be highly uncertain.

4. COST-EFFECTIVENESS

4.1. EAG comment on company's review of cost-effectiveness evidence

The company conducted a SLR of previous economic evaluations, the searches for which were considered well-structured and executed using a good range of sources. However, as stated in CS B.3.1, the company did not use the findings from their SLR to inform their economic model structure. The company applied a different filter to results from the same search to identify resource use and cost evidence but did not report in CS B.3.5 if or how the SLR was used to inform cost and resource use data selection and assumptions. Similarly, in CS B.3.4, the company reported conducting a SLR of health-related quality of life evidence, but it was not clear the extent to which and how the findings from this review informed the company's approach to patient utility assumptions, aside from a key mapping study being identified outside of the SLR, as noted in section 4.2.8 of this report. The EAG noted that in the company's HRQoL SLR, case reports and conference abstracts were excluded. It would be better practice not to use a 'study type' filter for these searches and to use a utilities filter instead; it was possible that some relevant data may have been missed as a result, if it was in a paper reporting a different type of study not included in the filter.

4.2. Summary and critique of company's submitted economic evaluation by the EAG

4.2.1. NICE reference case checklist

Table 17: NICE reference case checklist

Attribute	Reference case	EAG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	✓ No comment
Perspective on costs	NHS and PSS	✓ No comment
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	✓ No comment
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The company's lifetime horizon may be considered sufficient but excessive, in the context of a treatment that neither extends survival nor offers expected

		long-term health benefits after treatment cessation
Synthesis of evidence on health effects	Based on systematic review	The company reported conducting relevant SLRs but it was not clear how these reviews informed data selection and synthesis choices in the company's analysis
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	Health effects were expressed in QALYs. EQ-5D data were not collected in the TRuE-V trials. In section 4.2.8 the EAG explains and critiques the company's multi-step and multi-source approach to measure and value health effects
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	HRQoL data collected in the TRuE-V trials were not used in the company's multi-step and multi-source approach to measure and value health effects
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Relevant preference data were used within the company's multi-step and multi-source approach to measure and value health effects, as explained in section 4.2.8
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	✓ No comment
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	✓ No comment
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	✓ No comment

Key: EQ-5D, EuroQol 5 dimension; HRQoL: health-related quality of life; NHS, National Health Service; PSS, Pseronal Social Services; QALY: quality-adjusted life year; TA: technology appraisal

4.2.2. Validation

Upon receiving the company's model, internal checks were performed to ensure that the flow of patients and calculations behaved as intended. These included simple validity checks, the

assessment of cost and clinical inputs and review of Microsoft Excel® spreadsheet and Visual Basic for Applications® logic.

The submitted model passed standard internal consistency and stress checks performed by the EAG. However, sheet-by-sheet EAG review of the company model revealed clear issues with logic applied in the analysis, documented in section 4.2.5. Clinical inputs including patient characteristics, response status, the probability of treatment discontinuation and rates of adverse events were pooled across the TRuE-V studies. The EAG noted an error in the numerator of the 'no regain response' calculation, which related to miscategorising missing data entries as responses, covered in section 4.2.7.

Cost references were deemed suitable if taken from the most up-to-date sources relevant to the perspective of the National Health Service (NHS) and personal social services (PSS). Section 6.1 explains any amendments made to cost inputs provided by the company, using either drug costs within the company's provided appendix, or from the NHS Drugs and pharmaceutical electronic market information tool (eMIT) if less expensive.

4.2.3. Population

The company reported MHRA marketing authorisation for ruxolitinib to treat NSV with facial involvement in adults and adolescents aged 12 years and older, consistent with the TRuE-V1 and TRuE-V2 study populations informing the license. The final NICE scope listed no subgroups of interest. The British Association of Dermatology's professional organisation submission expressed a need for ruxolitinib early in the treatment pathway: "*Due to the lack of licensed treatments for vitiligo, and the fact that usually first line treatment for vitiligo includes topical preparations (TCS or TCI), ruxolitinib would fit into the first line treatment category alongside TCS and TCI and perhaps following a short trial of TCS or TCI*". Nevertheless, the company's economic analysis considered a subgroup of the licensed population, for whom a NICE recommendation is being sought: people aged 12 years and older for whom first-line topical treatments (TCS or TCI) are not suitable. Specifically, "*patients whose disease has not responded to TCS or TCI, or for whom TCS or TCI are contraindicated, not tolerated or otherwise medically inadvisable*". Clinical advice to the EAG was that this was a reasonable potential position for ruxolitinib, given ease of access to TCS and TCIs, and that a reasonable minority of people with vitiligo respond to these treatments. In clinical practice, the EAG's adviser noted that TCSs are typically tried first. If this does not work, tacrolimus (TCI) would be the next option to consider, or possibly TCS under occlusion. After exhausting topical treatment

options, NB-UVB and other second-line treatment options would be considered. The EAG's adviser considered that ruxolitinib could be used after other topical treatments had been exhausted, but before NB-UVB and other second-line treatment options were considered.

In the economic analysis, the company labelled the population of interest the "prior therapy" population. From the company's August 2023 original evidence submission, "prior therapy" did not appear to be a prespecified subgroup of the TRuE-V1 and TRuE-V2 trials, as discussed in Section 3. TCS or TCI exposure or suitability was a factor neither in the analysis populations tabulated in section 5.1 of the TRuE-V1 CSR, nor in the subgroups listed in section 9.5 of the same document. As such, the precise definition of the "prior therapy" subgroup whose data informed many elements of the economic analysis was not explicitly clear. What was more easily inferred was that the "prior therapy" subgroup did not cover all patients in the marketing authorisation; those who have not previously received therapy but for whom TCS or TCI are contraindicated or medically inadvisable are not represented. The EAG were unclear on the generalisability implications of this issue and noted it as an area of uncertainty for decision-making. This issue is captured by Key Issue 2.

"Prior therapy" subgroup data naïvely pooled across TRuE-V1 and TRuE-V2 samples informed baseline age, weight and gender characteristics in the economic analysis. These characteristics partially informed patient utility and treatment cost assumptions as described in 4.2.8 and 4.2.9. Pooled "prior therapy" TRuE-V1 and TRuE-V2 outcomes data informed treatment effectiveness and patient utility assumptions, as described in 4.2.7 and 4.2.8. Pooled TRuE-V1 and TRuE-V2 ITT data and estimates from the wider literature are used as proxy data in some instances, as noted throughout 4.2.7 to 4.2.9.

4.2.4. Interventions and comparators

The intervention in the company's analysis was ruxolitinib 1.5% cream, self-administered. The recommended dose is a thin layer of cream applied twice daily to the depigmented skin areas up to a maximum of 10% of BSA[1]. Clearly, the dose will vary by patient, based on varying extent of depigmentation and BSA across patients, varying interpretations of "thin layer" and "10% of BSA" across patients, and varying adherence to recommendations across patients. Based on clinical advice to the EAG, the dose used in practice may also vary depending on which areas of skin are considered the most important to patients. The SmPC and Information for patients leaflet each stated that no more than two 100g tubes per month should be used[1].

The company's analysis assumed that exactly 4.03g ruxolitinib was applied per day, which the company reported was the "*TRuE-V pooled median weight of study drug applied daily during 24-week period*", across ruxolitinib and vehicle cream arms. The EAG had several concerns with this approach to ruxolitinib dose calculation, all of which were agreed with clinical expert advice:

1. Overarching more specific concerns, uncertainty around the amount of ruxolitinib used in practice was important for expected cost-effectiveness results.
2. The EAG considered the use of vehicle cream dosing data in combination with ruxolitinib dosing data to estimate expected ruxolitinib dosing data to be inappropriate, when ruxolitinib dosing data could be used in isolation.
3. The EAG was concerned that in practice, with less medical oversight than in a trial setting, patients may be inclined to use more ruxolitinib, whether that means applying ruxolitinib more thickly or across more skin surface area. Patients may in practice have in mind the stated limit of two tubes a month. This equates to 6.57g (2dp) per day, [redacted]% (2dp) more than the daily dose assumed by the company.
4. The EAG was mindful that wastage; caused for example by accidently squeezing to excess, or by loss or mis-storage of the tube; may be more likely in practice than in a trial setting.
5. The EAG was conscious that patients would be issued 100mg tubes, and that any unused medicine in an open tube at the point of discontinuation would be wasted.

   Though not mentioned in the CS, the ruxolitinib tube sizes in TRuE-V studies were different to those that would be available in practice. In addition, maximum recommended use was higher in the TRuE-V studies than was advised in the UK label. The published protocol[32] for the TRuE-V studies stated that ruxolitinib was provided to patients in 60g tubes, and that participants were advised to limit use to no more than one 60g tube every week; 240g every 28 days.
6. The EAG was concerned that prescribing practice may tend towards the two 100mg tubes per month limit, even if patient use does not. For example, some patients may use less than two tubes each month but be prescribed two tubes per month nonetheless. The EAG's clinical adviser suggested that the company might consider producing smaller tubes. The EAG noted that the company produced 60g tubes for use in TRuE-V studies (point 5).

To partially address the first and second of these concerns, the EAG asked the company to provide further trial dosing data as a priority EAG question (B10). In response, the company stated that it was not possible to provide anonymised patient-level dosing data during the time available but did provide further summary data that shed further light on dosing differences across and within trial arms. These data are partially reproduced in Table 18, below.

Table 18: TRuE-V1[18] and TRuE-V2[19] ruxolitinib exposure summary statistics, adapted from Tables 6 and 7 of the company's response to EAG Clarification Q B10.

Variable	Ruxolitinib 1.5% Cream BID – TRuE-V1		Ruxolitinib 1.5% Cream BID – TRuE-V2	
Average weight of study drug applied daily	***During the double-blind period (g)***	***From Day 1 to Week 52 (g)***	***During the double-blind period (g)***	***From Day 1 to Week 52 (g)***
N	221	[redacted]	228	[redacted]
Mean (s.d.)	5.82 (16.587)	[redacted]	8.86 (31.385)	[redacted]
Median	4.17	[redacted]	3.96	[redacted]
Min, max	0.4, 237.1	[redacted]	0.4, 237.0	[redacted]

Abbreviations: BID, twice daily; s.d., standard deviation

Table 18 shows summary statistics for daily weight of ruxolitinib used across TRuE-V studies[18,19], stratified by study and timeframe from baseline. These data allow a focus on ruxolitinib data only, to estimate expected ruxolitinib doses, addressing the second of the EAG's initial concerns. In terms of the distribution of the data, the data were clearly skewed to the right: mean daily ruxolitinib use was notably higher than median daily ruxolitinib use, across studies, in the double-blind period [redacted]. The standard deviation and minimum and maximum statistics in Table 18 further clarify the distribution of drug use across the study samples. The data indicated that some TRuE-V patients used more ruxolitinib than was recommended in the license wording. As noted above, two 100mg tubes a month equates to 6.57g (2dp) per day; less than the mean daily use in TRuE-V2 in the double-blind period [redacted].

The implication of uncertainty around dosing in the TRuE-V studies[18,19] and of expected ruxolitinib use in practice upon cost-effectiveness estimated are explored in section 6 of this report, though the EAG's clinical expert has reassured the EAG to an extent on some of these potential issues. Without a strong position on expected ruxolitinib use in practice, the EAG's expert noted that it may be the case that patients become less adherent as time goes on, and the burden of two applications per day, alongside other skin applications many vitiligo patients

use, may lead to less use than recommended. On the other hand, as shown in Table 18, in TRuE-V1 and TRuE-V1, [redacted]

As described and critiqued in section 2.4 of this report, the company proposed vehicle cream, the control treatment in TRuE-V1 and TRuE-V2 trials, as the sole relevant comparator for their proposed positioning. As detailed in 2.4, the EAG was not convinced by this proposition. In short, vehicle cream was, by definition, not expected to have any effect in isolation. As such, vehicle cream was not part of the treatment pathway as described by BAD guidelines[21] and summarised in 2.3. Given the company's proposed positioning, EAG clarification question B1 (marked "Priority") asked the company to respecify the cost-effectiveness comparison at an appropriate point in the treatment pathway, with appropriate comparators. As documented in section 2.4 of this report, the company declined to do so.

Submissions from Vitiligo Support UK and BAD, supplemented by conversation with the EAG's clinical expert; an author of the BAD submission; have helped clarify the treatment landscape for the EAG, as discussed in Section 2. While the BAD treatment pathway in Figure 4 of the CS reflected the active treatment pathway for the relevant patient population, it became clear to the EAG that many vitiligo patients become lost to the system, owing primarily to system delays and the patient burden of some treatment options. In the first instance, GP prescription of topical first-line treatment may not be continued sufficiently long enough for a full treatment effect to manifest, which may lead to patient disengagement. Referral to secondary care is typically long and can involve a wait of up to a year. Once accessed, topical treatments (TCS and/or TCIs) may be tried again or for the first time under dermatologist direction, before NB-UVB is recommended for most patients (alone or in combination with topical treatments). However, NB-UVB is burdensome for the patient in requiring presentation at the secondary care centre two to three times a week. For some patients, for example adolescents in secondary education, this is not feasible. Patients can become disengaged at any point of the treatment pathway.

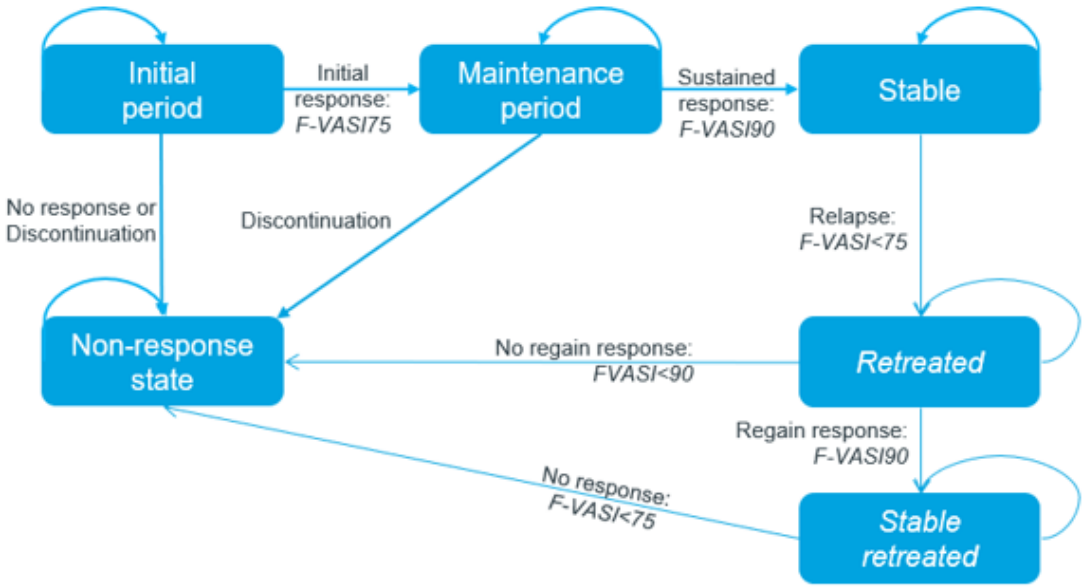
As such, it was the EAG's view that vehicle cream could be considered an appropriate proxy comparator for no treatment, at the end of the treatment pathway. This was inconsistent with the company's proposed positioning of ruxolitinib into the secondary care setting as an option after current 1st line treatment options (i.e., when other topical treatments have either been tried or are otherwise inappropriate). In this setting, the EAG's clinical advisor was clear; a dermatologist would try another option. The introduction of ruxolitinib here would displace,

delay, or add to the second-line BAD-recommended treatment options: NB-UVB with or without topical TCS or TCI, or for patients with progressive disease, betamethasone with NB-UVB.

4.2.5. Model structure and logic

The company's economic analysis comprised a *de novo* cohort-level model built in Microsoft Excel®; the company's model schematic is reproduced in Figure 5, below. Movements between states were allowed every 4 weeks. General population mortality data were used to capture the probability of death in each cycle; vitiligo was assumed to have no effect upon mortality, and no health state in the company's model was associated with a higher or lower chance of death than another.

Figure 5: Company's model structure schematic (CS Figure 20)



Note: Dead, not presented in the figure for simplicity, is an absorbing state and can be reached from any of the other health states

Identical model cohorts across (i) ruxolitinib and (ii) vehicle cream arms of the analysis entered the model in the "Initial period" state. Patients could discontinue into the "Non-response" state at the end of any of the initial model cycles, or otherwise either discontinue or continue into the "Maintenance period" state at the end of the "Initial period", based on whether or not F-VASI75 had been achieved. The CS reported the "Initial period" to end at 24 weeks, in line with the timing of the primary endpoint assessment in TRuE-V1[18] and TRuE-V2[19] However, this was not the case in the company's model, for two seemingly unintended reasons. First, the company

used a half-cycle correction, implying health state transitions at the mid-point of each 4-week cycle. Second, the initial period in the company's analysis lasted seven 4-weekly cycles, not six.

The EAG were mindful of clinical advice that F-VASI was not a measure used routinely in clinical practice, owing to its time-intensive nature. However, advice to EAG also noted that it is a robust and appropriate registrational trial measure, and that in aiming to capture repigmentation it is similar in its intentions to the more rudimentary measurements used in clinical practice. Perhaps more consequentially, the assumption that patients would discontinue owing to lack of efficacy if they have not achieved F-VASI75 at 24 weeks (the primary endpoint in TRuE-V1[18] and TRuE-V2[19]) was not in line with current NHS clinical practice. It was also not in line with the TRuE-V trials, in which all patients could receive ruxolitinib from week 24 to week 52, during the open-label extension period. The EAG's clinical advisor explained that they would assess a patient every 3-4 months and look for around 20% improvement at each visit to justify treatment continuation (i.e., a seemingly lower threshold of response is sought in practice, versus what is proposed by the company in its model). The SmPC states: "*Satisfactory repigmentation may require treatment beyond 24 weeks. If there is less than 25% repigmentation in treated areas at week 52, treatment discontinuation should be considered.*".[1] As such, in assuming all patients who have not achieved F-VASI75 at 24 weeks would discontinue to a non-response state, the company underestimated the proportion of patients who would continue treatment (and continue to accumulate treatment-related health benefits and costs) after 24 weeks if current practice for determining treatment continuation remains unchanged. This issue is compounded by the company's structural assumption that patients who enter the non-response state cannot move to another alive state. The non-response state is the state the company assumed was associated with the lowest patient utility and the second highest cost (after those in which ruxolitinib treatment costs are incurred), as described and critiqued in sections 4.2.8 and 4.2.9, respectively. Clinical advice to the EAG was clear: a consultant dermatologist would look to another treatment option after topical treatment.

After the "Initial period", the company assumed that those who were routed to maintain treatment remain in the "Maintenance period" state unless they discontinued treatment or died, until the cycle starting at week 56 (though the company reported this occurring at week 52, in line with the end of the TRuE-V open label extension period). The company's analysis assumed a time-invariant monthly discontinuation probability in the maintenance period that is distinct from that of the "Initial period". The data and assumptions informing the company's approach to discontinuation are described and critiqued in greater detail in section 4.2.7.

Though it is not represented in Figure 5, the company partitioned the “Maintenance period” state by response status; F-VASI75-89 versus F-VASI≥90; which allowed different utility assumptions to be applied for these two groups. However, this partitioning led the company to make structural assumptions that were unexplained in the CS and seemed to be unintentional and illogical. For example, it was structurally impossible for a patient in the “Maintenance period” state with F-VASI75-89 to achieve F-VASI≥90 and therefore transition to the “Stable” state; from “Maintenance period” state with F-VASI75-89, it was only possible to move to “No response” or death. Further, the company implied in the CS (B.3.2.2) that sustained response was defined by achieving F-VASI≥90 after the TRuE-V open-label extension phase, and that achieving sustained response should trigger moving onto the “Stable disease” state, where treatment was no longer needed. Yet, in the analysis, a cycle probability of sustained response (calculated from TRuE-V pooled data as described in section 4.2.7) was multiplied by the at-risk membership of the “Maintenance” F-VASI≥90 partitioned state. The company’s description of sustained response would suggest that the full at-risk membership of the “Maintenance” F-VASI≥90 partitioned state should transition to the “Stable disease” state for the next model cycle, after the TRuE-V open-label extension phase. In short, the EAG was concerned that the company had not modelled transitions from the “Maintenance period” to “Stable disease” as intended or stated.

In the “Stable disease” state, it was assumed that only disease management costs were incurred. There was no other state with higher assumed patient utility or lower assumed healthcare costs, as described and critiqued in sections 4.2.8 and 4.2.9. As illustrated in Figure 5, with the exception of moving to ‘dead’, it was only possible to move from “Stable disease” to “Retreated” (i.e., ultimately, all surviving patients would eventually move to re-treated provided the model time horizon was long enough). This occurred based on a time-invariant cycle probability of F-VASI <75, as detailed in section 4.2.7. It was an intentional structural limitation of the company’s model that movements from “Stable disease” to “Non-response” were not possible. The company assumed that retreatment was with the same topical treatment as used previously. On the intervention arm, this meant retreatment with ruxolitinib. Expert advice to the EAG suggested that retreatment with ruxolitinib would be rational, if near complete repigmentation was achieved, prompting discontinuation, which then led to depigmentation. On the comparator arm of the company’s analysis, this meant retreatment with vehicle cream. This was clearly not a reflection of clinical practice.

From “Retreated”, it was possible to transition to “Stable retreated”; which was equivalent to “Stable” in its cost and patient utility assumptions; or to the non-response state. The probability of transitioning to each of these states, and of transitioning to non-response from “Stable retreated”**,** was determined by F-VASI. It was assumed that there was no chance of discontinuing treatment for reasons other than the achievement of stability or loss of efficacy.

Eventually, the distribution of alive patients in the company’s analysis tended towards the non-response state, as indicated by Figure 5. In the company’s base case, this manifested as 95% of the cohort being in either “Non-response” or dead by around 8.5 years. As noted above, the non-response state was assumed to be associated with a high cost and low patient utility. The high assumed cost was driven primarily by the assumption that patients in this state incurred a monthly “Hospital-based NB-UVB” cost of £643.24 as described in section 4.2.9. Clinical advice to the EAG suggested that in practice, patients in long-term non-response would be likely to become disengaged with the healthcare system, with unmet need and low healthcare costs. It seemed that the company had received similar advice, but looked to reflect this in the model by assuming that all non-response health state costs discontinue after exactly 10 years: “*Disease management costs in the non-response health state are assumed to apply for the first 10 years only since the start of model simulation following input from clinical experts who stated that patients would consider discontinuing treatment and visits to the healthcare specialists after a certain period without any improvement*” (CS, p130). The company’s decision to select a 10-years-from-baseline time point, which was not linked to the duration of time any response to treatment was achieved, was not substantiated.

Overall, the EAG registered a range of concerns with the company’s model structure and logic, both with the company’s intended model design and its limited reflection of the EAG’s understanding of the vitiligo treatment pathway in NHS England practice, and, in places, with the rationale of logic in the company’s model, given the CS description of intentional design.

4.2.6. Perspective, time horizon and discounting

The perspective of the company’s analysis was that of the NHS and PSS on costs and that of patients on health effects, in line with the NICE reference case [33]. The company discounted cost and health outcomes at 3.5% per annum, also in line with the NICE reference case.

The perspective of the company’s analysis was lifetime. In the base case deterministic analysis, the mean age of the “prior therapy” subgroup (37.8 years) was assumed for the cohort at

baseline. The model's time horizon was set to 64 years, taking this cohort to age 101.8 years by the end of the time horizon. A lifetime horizon was sufficient but excessive, in the context of a treatment that neither extends survival nor offers expected long-term health benefits after treatment cessation.

Figure 6, produced by the EAG within the company's model, illustrates how the ICER produced by the company's list price deterministic base case analysis changed as the model's time horizon was varied from 5 to 64 years.[*] The figure shows that calculations beyond a 30-year time horizon had little impact upon the headline deterministic result. The figure also illustrates how the ICER fell as the time horizon increased from 5 years to 10 years, then increased at a decreasing rate as the time horizon increased beyond 10 years. The reason for the 10-year pivot in Figure 6 is the company assumption; discussed and critiqued in Section 4.2.5; that costs in the "non-response" state ceased to occur after exactly 10 years from model entry.

Figure 6: Relationship between model time horizon and company's list price base case deterministic ICER

Abbreviations: ICER, incremental cost-effectiveness ratio.

[*] To reduce the time horizon to 5 years, the EAG had to override data validation settings. This analysis was purely intended to demonstrate the relationship between the ICER and the time horizon in the company's model.

4.2.7. Treatment effectiveness and extrapolation

Broadly, the transition probability estimates governing movements between the model health states in Figure 5, as described and critiqued in Section 4.2.5, were based on a combination of summary TRuE-V1[18], TRuE-V2[19] and TRuE-LTE[9] data and assumptions. Table 33 of the CS was set out as summarising the key trial data and assumptions applied in the company's analysis, and is reproduced below as Table 19 for reference.

Table 19. Key TRuE-V1[18], TRuE-V2[19] and TRuE-LTE[9] data and assumptions applied in the company's analysis (CS Table 33)

Response category	Ruxolitinib cream		Vehicle cream		Section	Source
	Efficacy	SE	Efficacy	SE		
Initial and sustained response						
Initial response (F-VASI75 at week 24)*	[redacted]	[redacted]	[redacted]	[redacted]	Section B.3.3.3.1	Derived from pooled results of TRuE-V1 and TRuE-V2 data (Phase III) [26]
Sustained response F-VASI90 at week 52	[redacted]	[redacted]	Equal treatment effect assumed		Section B.3.3.3.2	
Relapse						
Time to relapse data (i.e., time to F-VASI<75) at week 104	Equal treatment effect assumed		[redacted]	[redacted]	Section B.3.3.3.3	Derived from Cohort A TRuE-V LTE (Phase III) [18]
Retreatment						
Regain response (F-VASI90 at week 104)	Equal treatment effect assumed		[redacted]	[redacted]	Section B.3.3.3.4	Derived from Cohort A TRuE-V LTE (Phase III) [18]
No regain response (F-VASI<75 at week 52 and F-VASI<90 at week 104)	[redacted]	[redacted]	Equal treatment effect assumed		Section B.3.3.3.4	Derived from Cohort B TRuE-V LTE (Phase III) [18]
Loss of response following retreatment (stable retreated)	Equal treatment effect assumed	[redacted]	Equal treatment effect assumed		Section B.3.3.3.4	Derived from Cohort A TRuE-V LTE (Phase III) [18]
Discontinuation						

Response category	Ruxolitinib cream		Vehicle cream		Section	Source
	Efficacy	SE	Efficacy	SE		
Initial period	[redacted]	[redacted]	[redacted]	[redacted]	Section B.3.3.6	Derived from pooled results of TRuE-V1 and TRuE-V2 data (Phase III) [26]
Maintenance period	[redacted]	[redacted]	[redacted]	[redacted]	Section B.3.3.6	

Notes: *Initial response is broken down into mutually exclusive FVASI75-89 and FVASI90 categories for modelling purposes. ** No regain response was calculated using the simple average of two approaches to missing data (removing missing data and treating missing data as non-responders). Data presented in this table has been derived from pooled results of the TRuE-V studies and/or TRuE-V LTE. These data are presented in Appendix M.

Abbreviations: F-VASI, facial vitiligo area scoring index; F-VASI75, 75% or greater improvement from baseline in F-VASI; F-VASI90, 90% or greater improvement from baseline in F-VASI; NR, Not reported; SE, standard error.

As described in Section 4.2.5, key problems with the application of TRuE-V data lay in the discord between efficacy endpoints in the regulatory studies and the EAG's understanding of effectiveness definitions in clinical practice. The data in Table 19, above, illustrate the scale of potential discord. For example, though it is not clear from Table 19, [redacted]% of the ruxolitinib arm of the pooled TRuE-V Prior Therapy subgroup achieved F-VASI75-89 and [redacted]% achieved F-VASI90 after 24-weeks' of treatment. All remaining ruxolitinib patients (100% - ([redacted]% + [redacted]%) = [redacted]%) were assumed to discontinue "Non-response" in perpetuity at the end of the model's 'Initial period', owing to lack of efficacy. This included [redacted]% of patients who achieved F-VASI50-74, as reported in the company's model. That is, patients who achieved a 50-74% F-VASI improvement after 24 weeks of ruxolitinib were assumed to be discontinued and consigned to interminable non-response. Given clinical advice received, the EAG considered that this may lack face validity.

The reporting in Table 19 was confusing and somewhat misleading for a number of reasons, but perhaps cardinally in its misreporting of the data presented. For example, the entries in the "Loss of response following retreatment" row of Table 19 are clearly erroneous in places, indicative of a copy-paste error.

Elsewhere, what Table 19 shows as a "No regain response" estimate was applied in the analysis as a 4-week (model cycle) probability estimate. For this input, the way the company accounted for missing data in their calculations was erroneous. On p120 of the CS, the company wrote:

"Two methods to account for these missing data were used in the analysis: firstly, removing missing data from the overall sample of those with F-VASI<75 at week 52 (n=99) and secondly, treating missing data as non-responders.

"For the first method, the probability of having F-VASI<90 at week 104 if patients with F-VASI<75 at week 52 is calculated as [redacted]% ([redacted]); for the second method the probability is [redacted]% ([redacted]). In the base case, a simple average of the two methods is applied, giving an overall probability of [redacted]%."

The first method assumed that non-responders were missing at random, a strong assumption applied without sufficient evidence, but applied correctly. The second method did not treat missing data as non-response data, as described by the company. Instead, the approach miscategorised the n=[redacted] missing data entries as responses. If instead these entries had been categorised as non-response entries, the probability of non-response would have been calculated as ([redacted]+[redacted]) / [redacted] = [redacted] / [redacted] = [redacted]% (2dp). The EAG considered it more appropriate to assume that missing data were indicative of non-response than missing-at-random, and favoured this choice over the simple average of the two favoured by the company.

Part of the simplicity of the company's approach to treatment effectiveness and extrapolation was the time-invariant nature of transition probability assumptions applied. In a cohort-level Markovian model such as the company's, applying time-varying probabilities can be cumbersome, but applying time-varying probabilities that vary only from baseline is not. The company's approach to time to discontinuation assumptions, using data shown in Table 19, was a clear example of assuming time-invariance when assuming time variance may be appropriate given the data, meaningful for results, and uncomplicated to apply in the economic analysis. As such the EAG requested TRuE-V1 and TRuE-V2 Kaplan-Meier time to treatment discontinuation data, stratified by study and treatment arm, for relevant TRuE-V1 and TRuE-V2 populations, as part of EAG clarification question B8. In response, the company provided Kaplan-Meier data as requested, though without reporting censor points. A summary image of treatment discontinuation projections for the "Prior Therapy" population provided as part of this response is reproduced as Figure 7, below.

In a separate part of clarification question B8, the EAG also requested that the company use these data to incorporate functionality into the revised model to allow time-to-treatment discontinuation to be accurately modelled. The company did not do this. In light of the ruxolitinib data in Figure 7, the EAG were not overly concerned by the assumption of time-invariant

treatment discontinuation in the analysis, though more accurate use of the available treatment discontinuation data would have been preferred in the first instance. What was of more concern was how closely the company's model reflected the use of ruxolitinib in the TRuE-V open-label extension periods. As indicated in Figure 7, around 80% of patients randomised to ruxolitinib were still receiving ruxolitinib a year into treatment, and six months into the open-label extension. In the company's analysis, less than 25% of patients in the ruxolitinib arm were modelled as remaining on maintenance treatment at 1 year. Clinical advice to the EAG was noncommittal on expected length of treatment, beyond the expectation of noticeable improvements every 3-4 months justifying treatment continuation and anecdotal consideration that the burden of treatment application may take its toll months and years into treatment. In short, with respect to treatment discontinuation assumptions, the company's analysis was neither reflective of evidence from its own registrational studies nor expected clinical practice.

Figure 7: Kaplan-Meier treatment discontinuation projections stratified by TRuE study and treatment arm, as presented by the company in response to EAG clarification question B8.

Overall, the EAG considered the company's methods to incorporate data from the TRuE-V studies into the model to be subject to a number of substantial limitations. Ultimately, this meant that the EAG had little confidence in the results of the model. These issues were an overarching concern to the EAG, and in combination with concerns with the company's model structure and logic detailed in Section 4.2.5, comprised Key Issue 3. Until resolved, this prevented the EAG from presenting more than a tentative preferred base case, as discussed in Section 6.

4.2.8. Health-related quality of life

Within the TRuE-V studies, EQ-5D data were not collected. Data from other health-related quality of life measures were collected, including DLQI and VitiQoL instruments, as described in Section 3.2.2.5. Outside of the TRuE-V[18,19] studies, the company identified 24 studies investigating the HRQoL of people with vitiligo in a SLR of HRQoL evidence. However, the company did not use any TRuE-V DLQI data to inform utility assumptions in their model, nor any VitiQoL data collected beyond baseline. Furthermore, the company did not use data from any of the 24 studies identified in their SLR.

Instead, the company used an opaque and loosely justified approach to derive health state utility values that involved using F-VASI results from TRuE-V1 and TRuE-V2 in combination with a published mapping algorithm and various assumptions. To understand exactly how these utility values were generated, the EAG interpreted company reporting across (i) the appropriate section of the CS, (ii) Appendix O of the CS, (iii) a technical report embedded within Appendix O of the CS, (iv) Appendix I of the technical report embedded within Appendix O of the CS, and (v) protocols and Excel files containing regression analysis results embedded within Appendix I of this technical report. Following this, the EAG confirmed the accuracy of the EAG's interpretation of the company's multistep approach in clarification question B13. The confirmed company approach and its implicit assumptions are as follows.

1. A mapping study provided a means of generating UK vitiligo patient EQ-5D-5L utility values from RPS, VNS and VitiQoL data was identified outside of the company's systematic review[30]. In this study, Begum et al. estimated mapping algorithms using data on this range of outcomes from the HI-Light study[3]. Begum et al. also reported algorithms allowing prediction of EQ-5D-3L utility from RPS, VNS and VitiQoL data, using the Hernandez et al[34] crosswalk recommended in the NICE Manual[35].

7. Assuming RPS score was a suitable proxy for F-VASI score, the Begum et al.-reported algorithm to predict EQ-5D-3L utility from RPS data was used in combination with patient-

level F-VASI data from the Prior Therapy TRuE-V sample to generate post-baseline proxy patient-level EQ-5D-3L utility estimates for said TRuE-V patient sample.

That is, the following equation as reported by Begum et al.

$EQ\text{-}5D_{RPS} = 0.709 + (0.0119 * \mathbf{RPS}) - (0.000214 * \mathbf{RPS}^2) + (0.00000118 * \mathbf{RPS}^3)$

was interpreted as equivalent to the following as reported in B.3.4.3 of the CS, where F-VASI Category and RPS are taken as interchangeable:

$EQ\text{-}5D_{F\text{-}VASI} = 0.709 + (0.0119 * \mathbf{F\text{-}VASI\ Category}) - (0.000214 * \mathbf{F\text{-}VASI\ Category}^2) + (0.00000118 * \mathbf{F\text{-}VASI\ Category}^3)$

8. As both F-VASI and RPS are measure of changes in pigmentation from baseline, such scores were not available at baseline. As such, baseline patient-level utility estimates were generated by applying baseline "prior therapy" TRuE-V sample VitiQoL scores to the following Begum et al. VitiQoL algorithm:

   $EQ\text{-}5D_{VitiQoL} = 0.9652 - 0.00205 * \text{Total VitiQoL Score}$

9. Next, the patient-level data utility data generated through steps 2 and 3 above were added to the TRuE-V dataset as if they were additional data fields, and regression analyses were performed to estimate the determinants of changes in patient utility from baseline to 24 weeks (and in a separate analysis not used in the company base case, of changes in patient utility from baseline to 52 weeks). The technical report embedded within Appendix O of the CS reported that a model of the following general form was applied to a stepwise selection procedure to determine final variable selection according to minimum Schwarz Bayesian Information criterion:

   $\text{Change from Baseline (CFB)} = \text{Baseline EQ-5D utility} + \text{Age} + \text{Sex} + \text{Skin Type (Fitzpatrick Scale)} + \text{Disease Status} + \text{Treatment} + \text{F-VASI 50 Response} + \text{F-VASI 75 Response} + \text{F-VASI 90 Response} + \text{Baseline EQ-5D} * \text{Treatment Interaction}$

   After final model selection, the technical report stated that predictions of CFB utility were derived through least squares means (marginal means) analysis on the final models after variable selection.

From this process, applied to the "prior therapy" TRuE-V sample, the company derived most of the estimates presented as the utility data informing the company analysis in Table 37 of the

CS, reproduced as Table 20, below. That is, the "No response", "F-VASI50-74", "F-VASI75-89" and "F-VASI90" estimates in Table 20 were generated through this process†.

Table 20: Reproduction of CS Table 37 – combined results informing the company's base case utility assumptions

State	Utility value: mean (standard error)	95% CI (Lower, Upper)	Justification
Baseline	0.879 (0.003)	0.874, 0.884	VitiQoL baseline utilised as F-VASI mapping produced no available baseline data [36]
No response	-0.082*	-0.087, -0.077	F-VASI (DP: -37.5%) was the best performing measure in the mapping algorithm [36]
F-VASI50-74	0.010*	-0.007, 0.028	
F-VASI75-89	0.056*	0.037, 0.074	
F-VASI90	0.066*	0.047, 0.084	

Abbreviations: F-VASI, facial vitiligo area scoring index; F-VASI50-74, 50% to 74% improvement from baseline in F-VASI; F-VASI75-89, 75% to 89% improvement from baseline in F-VASI; F-VASI90, 90% or greater improvement from baseline in F-VASI; VitiQoL, vitiligo-specific quality-of-life instrument
Source: Information presented in Section B.3.4.3.3 of the company's submission. Source data from Incyte, technical report for statistical analysis and utility modelling [Data on file] [36]

The data in Table 20 warrant careful interpretation that is lacking in the CS. First, the data presented under the column headed "Utility value: mean (standard error)" are not a collection of mean utility values as implied by the heading of the Table's second column. Instead, they are a baseline utility value and a collection of decrements applied separately to said utility value to generate health state utility values for the company's model. Table 21 summarises the expected utility values implied by the estimates in Table 20 and applied as health state utility values in the company's base case deterministic analysis. However, the outcome descriptions in the first column of Table 20 and Table 21 are different to the company's health state descriptions, as presented in the model schematic in Figure 5. Table 22 summarises the utility values associated with each model health state in the company's deterministic base case analysis.

† Note: The EAG highlights that the technical report embedded within Appendix O of the CS did not contain the results the EAG provides in Table 22.

Table 21: Absolute expected utility values implied by Table 18 and applied in the company's model

Description	Utility value assumed in company's deterministic analysis
Baseline	0.879
No response	0.797
F-VASI50-74	0.890
F-VASI75-89	0.935
F-VASI90	0.945

Table 22: Health state utility values applied in the company's deterministic analysis

Health state	Utility value assumed in company's deterministic analysis
Initial period	0.879
Maintenance period	0.935-0.945, dependent on response level
Stable	0.945
Retreated	0.879
Stable retreated	0.945
Non-response	0.797

There are clear issues with the company's approach to estimating and applying health state utility assumptions. Perhaps the most notable are listed as follows.

1. The number and strength of assumptions required to go from TRuE-V F-VASI data to the the utility values in Table 22, as set out in this section so far, called into question the reliability of the utility values as evidence-based estimates for decision making.
2. The quality of reporting by the company with respect to the approach and justification for each choice and assumption required to reach the utility values in Table 22 was a barrier to review and further reduced confidence in the appropriateness of the values selected.
3. As illustrated by EAG clarification question B15, the expected utility values assumed for "Maintenance period", "Stable" and "Stable retreated" health states were higher than the age-equivalent general population utility value from a source commonly cited in NICE appraisals[37]. Notably, this same source was used in the company's own model to adjust utility for the effect of ageing over the model's time horizon.

4. Importantly, all estimates other than that for the “Initial period” in Table 22 (“Baseline” in Table 21) were based on analysis of changes from baseline, where the baseline estimate was derived from a different measure and algorithm than all post-baseline data. Any interpretation of these values warrants extreme caution. When such values are in excess of general population estimates, there is clear reason to doubt their plausibility.
5. The company’s own approach estimated a utility value of 0.890 for patients achieving F-VASI50-74 at 24 weeks, as documented in Table 21. Yet, in the company’s model, patients achieving F-VASI50-74 are categorised as “Non-responders” and assigned a utility value of 0.797.
6. As noted in Section 3, the company’s clinical evidence submission did not demonstrate a treatment effect upon patient HRQoL, nor on important domains of HRQoL expected to be affected by ruxolitinib, such as anxiety and depression. The EAG was concerned that this evidence had been selectively set aside in preference of an approach that estimated a utility benefit from an F-VASI benefit.
7. As also noted in Section 3, the company’s evidence submission was clear in noting the humanistic burden of vitiligo. For example, the company cited that 54.2% of patients in the VALIANT study reported symptoms of moderate-to-severe depression in B.1.3.1.3 of the CS. Meaningful accounts of disease burden were also provided in Patient Body and Professional Organisation submissions from Vitiligo Support UK and the British Association of Dermatologists, respectively. The EAG was concerned that the company’s utility values may lack face validity in this context, in comprising a baseline utility value similar to an age-equivalent general population estimate and response-defined utility values that exceed this general population estimate.

For these reasons, the EAG noted the company’s approach to capture patient HRQoL effects in the cost-effectiveness analysis as a Key Issue (Key Issue 6). The importance of uncertainty around the company’s utility assumptions for cost-effectiveness results is explored in Section 6 of this report.

On top of the issues with the approach to capture health state utility values discussed above, the company’s economic analysis did not in any way account for the HRQoL implications of adverse events, despite treatment-emergent adverse events affecting 47.7% of ruxolitinib patients in the pooled TRuE-V population, as documented in Table 23 of the CS. In EAG clarification question B16, the EAG asked the company to incorporate into its analysis utility

implications for the adverse event data in Table 24 of the Document B of the company submission (treatment-emergent adverse events occurring in ≥ 1% of patients in any treatment group). In response, the company declined to do so, without reasonable justification. In the context of the potential magnitude of QALY gains for ruxolitinib, the EAG considered that it was plausible that accounting for the HRQoL implications of adverse events appropriately could meaningfully affect cost-effectiveness results. Accordingly, and as documented in Section 1.5, the absence of expected utility implications of adverse events from the company's analysis is, alongside issues with the company's approach to account for cost implications of such adverse events (documented in Section 4.2.9), EAG Key Issue 7.

4.2.9. Resources and costs

The company report identifying 11 publications reporting healthcare resource and cost data for people with vitiligo in their SLR of economic literature, but do not report if, how or for what reason any of these 11 studies were used to inform their cost assumptions. The company's cost assumptions broadly fell under three categories, defined here as: treatment acquisition and administration; disease management; and adverse event management. The company's approach to assume costs for each of these categories is documented and critiqued in this section in turn.

Issues with the company's approach to treatment acquisition and administration and disease management costs, documented across sections 4.2.9.1 and 4.2.9.2, led the EAG to conclude that the cost assumptions in the company's analysis biased cost-effectiveness results in favour of ruxolitinib. Together, these issues comprised EAG Key Issue 5. Issues with the company's approach to capture adverse event costs, documented in section 4.2.9.3, along with the absence of HRQoL implications of adverse events critiqued in section 4.2.8, together comprise EAG Key Issue 7.

4.2.9.1. Treatment acquisition and administration

The price of ruxolitinib was £[redacted] per 100g tube in the company's analysis. The company described this price as a list price [redacted]. The company assumed that the combined acquisition, dispensation and administration cost of ruxolitinib was equal to £[redacted] per 100g tube, at list price. That is, the company assumed that there was no dispensation or administration cost associated with ruxolitinib, which underestimated costs on the intervention arm of the analysis, particularly if ruxolitinib was considered as a potential end-of-line treatment, which the EAG considered to be a possibility. In

the CS, the company reported submitting a Patient Access Scheme (PAS) including a simple [redacted]% discount to the acquisition price of ruxolitinib. The company's cost-effectiveness results assumed that this PAS discount would hold in practice.

The PAS-adjusted ruxolitinib acquisition price combined with dosing and time-on-treatment assumptions comprised total ruxolitinib treatment costs in the company's analysis. Adjusted for PAS discount and the company's dosing expectations, the per 4-week cycle ruxolitinib acquisition cost in the company's analysis was £[redacted]. The company's approach to model ruxolitinib dosing was EAG Key Issue 4, as set out in Section 4.2.4, while the EAG also held concerns about the model's deviation from observed time-on-treatment in the TRuE-V study programme, as discussed in Section 4.2.7.

On the comparator arm of the company's analysis, the acquisition cost of vehicle cream was assumed to be equal to the per ml cost of suncream, using a British National formulary (BNF) estimate of £9.70 per 125ml bottle. Vehicle cream dosing was assumed to be equal to ruxolitinib dosing using pooled TRuE-V data, as described in in Section 4.2.4, assuming 1ml vehicle cream equals 1g vehicle cream. Adjusted for the company's dosing expectations, the per cycle vehicle cream acquisition cost in the company's analysis was £[redacted]. Vehicle cream time-on-treatment assumptions were based on TRuE-V vehicle cream time-on-treatment data in a similar manner to the ruxolitinib arm, as described in Section 4.2.7.

Concomitant treatments were assumed for all alive model states except for "Stable" states, where no treatment costs were assumed. For "initial", "maintenance" and "retreated" states, these costs were assumed equivalent across arms and totalled £17.66 per 4-week model cycle. Concomitant treatments were assumed to comprise suncream, vitamin D supplement, camouflage cream and fixing powder. The EAG noted that the company was therefore effectively applying a cost of suncream twice in the vehicle cream arm of the analysis. In addition, the EAG noted that in practice, ruxolitinib use would limit a person's freedom to apply suncream when wanted, as discussed in Section 3.2.2. The company did not account for this in their analysis.

Importantly, as well as intervention and comparator treatment costs, the company separately defined "BSC" treatment costs, which they assumed for the "non-response" state of each arm of the model. As patients on the vehicle cream arm of the analysis spent longer incurring costs in the "non-response" state over the model's time horizon, overestimating costs for this state would bias the analysis in favour of ruxolitinib. BSC treatment costs were assumed to comprise £19.05

of concomitant treatment costs (assuming different levels of concomitant treatment use than in "initial", "maintenance" and "retreated" states) and, as noted in Section 4.2.5, £643.24 of hospital-based NB-UVB costs, every 4-week cycle.

The EAG had several notable issues with the NB-UVB assumptions applied in the "non response" health state of the model:

1. The company assumed that [redacted]% of those in the "non-response" state received NB-UVB, based on a simple average of the proportion of UK patients who had ever used light, laser or NB-UVB therapy and the proportion of UK healthcare professionals who recommended such therapy in the Vitiligo and Life Impact Among International Communities (VALIANT) study; a global survey study exploring the natural history and management of vitiligo from patient and healthcare professionals[38]. In the EAG's view, such a source and consideration of an active, effective treatment assumption was inappropriate for a health state that was characterised as a "non-response" state. As noted in Section 4.2.5, clinical advice to the EAG suggested that in practice, patients in long-term non-response were likely to become disengaged with the healthcare system, with unmet need and low healthcare costs.
2. The company assigned a cost of £140.84 to every NB-UVB session, based on the 2021/22 NHS Reference Cost of an outpatient dermatology procedure. Assuming three sessions per week over nine months, the company assumed that this cost was incurred 117 times per course of NB-UVB; a total cost of £16,478.36. Yet, clinical advice received by the company, provided to the EAG in response to EAG priority clarification question B4, informed the company that for facial vitiligo, a home-based approach would be suggested, with handheld device training being provided in the hospital and use monitored every 3-4 months unless issues arose. This advice appeared to be in line with practice in the UK NHS-based HI-Light trial. A 2021 economic evaluation of topical corticosteroid and home-based NB-UVB based on HI-Light trial results used similar assumptions, and estimated a total NHS cost of a course of NB-UVB of £775[39]. Clinical advice to the EAG noted that hand-held NB-UVB devices were not available at every NHS centre; the average NHS cost of a course of NB-UVB may be greater than £775 as a result, but the EAG expected that £16,478.36 was a marked overestimation. Further, the NHS resource burden implicit in the company's costing assumptions would surely reduce the proportion of patients expected to be able to access NB-UVB treatment.

3. The company assumed that an equivalent of a nine-month course of hospital-based NB-UVB occurred every year, for those [redacted]% of patients in the "non-response" state assumed to receive NB-UVB. The £643.24 per cycle cost was calculated as [redacted]% * £140.84 * 177 sessions * (28 days / 365.25 days), and was assumed to apply to the "non-response" state every cycle, for as long as treatment costs were assumed plausible in this state – until 10 years from baseline in the company's analysis, as critiqued in Section 4.2.5. From the EAG's understanding of capacity constraints in NHS Dermatology departments, near continuous NB-UVB was not plausible for any vitiligo patient, let alone on average. The EAG's clinical advisor explained that there was no limit on the number of NB-UVB courses a patient could receive, but the decision to recommend a second course would be based on response to NB-UVB previously; even if retreatment with NB-UVB was recommended, there may be a one-year wait between NB-UVB courses. Further, there was a tendency within the NHS to prioritise NB-UVB capacity for patients with conditions that respond to NB-UVB more quickly than vitiligo.

With respect to EAG concerns over the company's proposed positioning of ruxolitinib (Key Issue 1), the EAG considered that a comparison to vehicle cream was only potentially appropriate as an end-of-line comparison. In this instance, assuming any NB-UVB use after ruxolitinib or standard of care treatment would be inappropriate. However, even if the company's positioning of NB-UVB after ruxolitinib treatment could be considered appropriate, the EAG identified clear issues with the company's characterisation of NB-UVB in the NHS setting. Overall, the company's approach to cost for NB-UVB treatment overestimated the expected cost in several ways, biasing cost-effectiveness results in favour of ruxolitinib.

4.2.9.2. Disease management

The company's analysis assumed that disease management costs were incurred every cycle, in each of the alive health states of the model. The categories of disease management resource that the company considered were: Dermatologist outpatient consultation; Dermatologist telephone consultation; Dermatologist nurse visit; GP consultation; Accident & Emergency (A&E) Visit; Psychological support. The company assumed that the amount of these resources used differed across (i) "Initial", "Maintenance" and "Retreated" states, (ii) "Stable" states and (iii) the "non-response" state. Generally, "Stable" states were assumed to use less disease management resource than "Initial", "Maintenance" and "Retreated" states, while the "non-response" state was associated with the highest resource use burden, as documented in Table 41 of the CS. The company reported that their resource use assumptions were based on a

combination of the 2021 Sachs et al. economic evaluation of HI-Light trial outcomes,[39] referenced in Section 4.2.9.1, and clinical expert opinion.

The company's disease management resource use assumptions implied engagement with the health service at least every 2 months, across alive health state, and produced the following 4-weekly disease management cost estimates:

- "Initial", "Maintenance" and "Retreated" states: £308.31
- "Stable" states: £132.37
- "Non-response" state: £548.22

The EAG was concerned that the company's approach overestimated disease management costs, in a manner that biased cost-effectiveness results in favour of ruxolitinib. Clinical advice to the EAG suggested that the company's psychological support assumptions were inaccurate. The EAG's clinical adviser noted that many clinicians did not screen for psychological distress; their expectation is that only around 15% of patients were directed towards psychological support resources. Even if this happens, the direction would be towards self-referral for NHS Talking Therapies (formerly IAPT: Improving Access to Psychological Therapies)[40] for those with moderate or severe distress and to self-help resources for those with mild distress. The EAG's clinical adviser's understanding was that waiting lists for such services meant that some self-referred patients disengage, while others may seek private pyschological support instead.

The EAG was also concerned with the assumption of ongoing Dermatologist appointments in the "non-response" state. As noted in Section 4.2.9.1 and elsewhere in this document, if a comparison to vehicle cream was only potentially appropriate for end-of-line positioning, no NHS Dermatology appointments would be expected in this "non-response" state. Even taking the company's proposed positioning, a per patient expectation of ongoing engagment with NHS Dermatologists around every 2 months for 10 years post baseline does not tally with the EAG's understanding of resource constraints in Dermatology departments and the impact of this on typical uptake.

Further, the EAG noted that the Dermatologist and GP appointment frequency assumptions for the "non-response state" were estimates from the NB-UVB + TCS arm of Sach et al[39]. As such, in the company's base case, the "non-response" health state costs included (i) NB-UVB treatment costs that were calculated based on the cost of Dermatology appointments, as

described in Section 4.2.9.1 and (ii) separately, Dermatology appointment costs to capture disease management resource use. This was clearly double counting.

Overall, for reasons documented here, the EAG was concerned that the company's analysis overestimated disease management costs in a manner that biased cost effectiveness results in favour of ruxolitinib.

4.2.9.3. Adverse event management

The company's analysis captured treatment-arm specific expectations for adverse event costs, using incidence rates of adverse events occurring in ≥4% of patients in either arm up to week 24 across TRuE-V1 and TRuE-V2, as reported across B.3.3.5 and B.3.5.3 of the CS (Doc B). This accounted for incidences of application site acne, application site pruritis, nasopharingitis, headache and upper respiratory tract infection; though the incidence of the latter appeared to have been <4% across arms, from Section B.3.3.5. The company's approach produced expected per cycle adverse event costs of £4.11 and £1.67, for ruxolitinib and vehicle cream arms of the company's analysis, respectively.

In EAG clarification question B16, the EAG asked the company to extend the scope of adverse events included in the cost calculation, to capture treatment-emergent adverse events occurring in ≥ 1% of patients in any treatment group. In reply, the company declined to amend their original approach. Although the technology under appraisal was a topical treatment and there were no clear safety concerns in the TRuE-V studies, the EAG asked for this, alongside consideration of the HRQoL consequences of such events, for several reasons. Firstly, 4% is an arbitrary and high cut-off, while 1% is an established cut-off for "common" adverse events, as noted in European Medicines Agency documentation[41]. Secondly, if this appraisal led to a positive recommendation for ruxolitinib at the end of the existing treatment line, it will replace no treatment, and thus definitively introduce toxicity. Thirdly, the dosing data received in response to EAG clarification questions and documented later in Section 6.2.3 suggested that some patients in the TRuE-V trials exposed themselves to more ruxolitinib than recommended, which may have resulted in safety issues unanticipated with intended use.

5. COST-EFFECTIVENESS RESULTS

5.1. Company's cost-effectiveness results

The company results presented throughout section 5 reflect the proposed PAS-adjusted price of ruxolitinib, as reported in section 4.2.9.1.

5.1.1. Base case results

The company's post-clarification questions deterministic and mean probabilistic base case cost-effectiveness results are summarised in Table 23. The deterministic analysis underestimated total QALYs and overestimated total costs across model arms, relative to the mean probabilistic analysis. The company's mean probabilistic results were produced using 2,000 probabilistic model iterations, with evidence of testing for robustness of summary results to additional iterations up to 2,000 iterations presented.

The EAG placed little weight on the company's summary base case results, owing to the various issues in the company's analysis documented through sections 0, 2 and 4. As documented in Sections 2 and 4 and comprising Key Issue 1, in the EAG's view the company's analysis did not address the decision problem, as defined by the company for a subgroup of the final scope. As a result, the analysis results were fundamentally of little value to the appraisal. Further issues identified throughout Section 4 together suggested bias in the company's analysis in favour of ruxolitinib.

Table 23: Company base case results

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained
Company deterministic base case					
Vehicle cream	[redacted]	[redacted]	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	£13,634
Company mean probabilistic base case					
Vehicle cream	[redacted]	[redacted]	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	£14,676

Abbreviations: QALYs, quality adjusted life years

5.2. Company's sensitivity analyses

Though the EAG placed little weight on the company's headline cost-effectiveness results, the company's sensitivity analyses had merit in characterising some of the parameter uncertainty around the company's results, and in illustrating some important areas of sensitivity in the company's analysis.

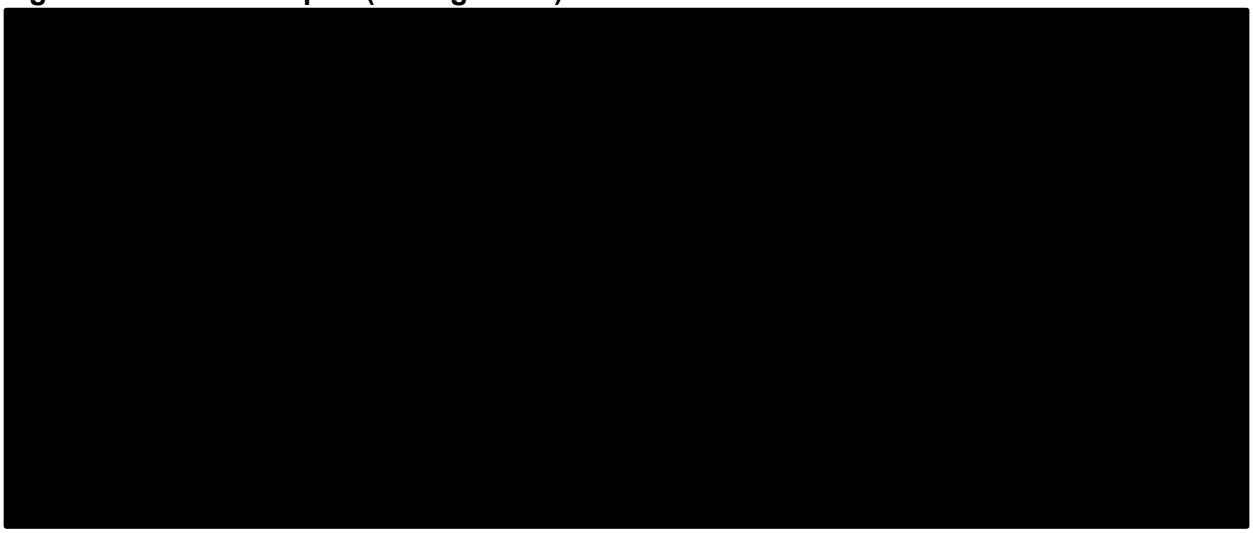
5.2.1. One-way sensitivity analysis

The company's one-way sensitivity analysis (OWSA) used a mixture of distributional assumptions across parameters, as partially reported in section B.3.6.2 of the CS. Section B.3.8.2 of the CS reported OWSA results as a tornado diagram showing the 20 parameters that led to the greatest variation in ICER results. Limitations in the parameter testing descriptions in this diagram meant that its reproduction in this report would be of little value. In short, base case deterministic company cost-effectiveness results were most sensitive to uncertainty around response rate estimates, discontinuation rate estimates and NB-UVB assumptions in the "non-response" state. In the extreme, the company's OWSA caused ICER estimates to vary from around £1,000 to over £30,000 per QALY gained.

5.2.2. Probabilistic sensitivity analysis

Figure 8, below, reproduces the probabilistic sensitivity analysis (PSA) scatterplot presented as Figure 22 in the CS. The distribution of PSA iteration results gave a picture of the parameter uncertainty around the mean PSA results in Table 23, above. The probability that ruxolitinib was cost-effective at a willingness-to-pay threshold of £20,000 per QALY gained in the company's probabilistic analysis was [redacted]%.

Figure 8: PSA Scatterplot (CS Figure 22)



5.2.3. Scenario analyses

The company presented results from seven scenario analyses. Descriptions of each of these scenarios, as reported by the company, are provided in Table 24. Scenario analysis results as reported by the company are shown in Table 25. Despite only comprising seven scenarios, the company's scenario analysis highlighted some important model sensitivity. Perhaps most notably, using the published mapping algorithm for the Vitiligo Noticeability Scale in place of the F-VASI algorithm from the same study (as described in section 4.2.8) caused the estimated QALY gain associated with ruxolitinib to diminish to [redacted] with a resultant ICER of £398,929. Another notable scenario illustrated the sensitivity of company-preferred results to when costs in the "non-response" state were no longer asumed to be incurred. When this parameter was set to 5 years from baseline rather than 10 years from baseline, the estimated incremental cost of ruxolitinib more than doubled and the estimated ICER increased to £39,272.

Two of the company's seven scenarios were analyses for different patient populations: (i) the overall TRuE-V population and (ii) the TRuE-V subgroup with Fitzpatrick skin type IV-VI. Using response and treatment discontinuation rates estimated from these samples and otherwise keeping all model settings constant, the company's analysis predicted that ruxolitinib was less cost-effective in the overall population than in the "prior therapy" population and highly cost effective (dominant) in the Fitzpatrick skin type IV-VI population.

Table 24: Overview of company's scenario analyses (CS Table 46)

No	Model scenario	Base Case	Description/Justification
1	Utility data source: F-VASI (DP: -25%)	Utility data source: F-VASI (DP: -37.5%)	This scenario explored the impact of utilising alternative bandings in the F-VASI mapping algorithm. Depigmentation categorisation I: Percentage change of -25% (i.e., all patients with depigmentation were truncated to having skin pigmentation loss not greater than 25%) [DP: -25%] (Section 4.2.8)
2	Utility data source: VNS	Utility data source: F-VASI (DP: -37.5%)	VNS was the secondary endpoint in the TRuE-V studies [32]. As such, mapping from this endpoint was considered for scenario analyses.
3	Model time horizon: 10 years	Model time horizon: Lifetime (63 years)	This scenario explores the impact of a shorter time horizon in the model.
4	Stop costs in the non-response state: 5 years	Stop costs in the non-response state: 10 years	This scenario explores the impact of varying the length of time costs are incurred in the non-response state. This aligns with clinical feedback where clinicians noted that patients experience treatment fatigue and patient choice varies over time [5].
5	Stop costs in the non-response state: Lifetime	Stop costs in the non-response state: 10 years	This scenario explores the impact of varying the length of time costs are incurred in the non-response state. This aligns with clinical feedback where clinicians noted that patients experience treatment fatigue and patient choice varies over time [5].
6	Overall population	Prior therapy sub-group	This scenario explores the impact of assessing the overall population recruited in the TRuE-V studies.
7	Patients from overall population with Fitzpatrick skin type IV-VI	Prior therapy sub-group	The Fitzpatrick IV-VI categorisation was chosen as darker skin types are associated with a greater patient burden [42] including use of a significantly greater number of treatments [43]. This categorisation has been used in a recent study which assessed the importance of facial involvement for patients [44].

Abbreviations: DP, depigmentation; F-VASI, facial vitiligo area scoring index; VNS, vitiligo noticeability scale

Table 25: Summary of key cost-effectiveness results from scenario analyses (CS Table 47)

No	Model scenario	Treatment	Total costs (£)	Total QALYs	Incremental Costs (£)	Incremental QALYs	ICER vs vehicle cream
1	Utility data source: F-VASI (DP: -25%)	Vehicle cream	[redacted]	[redacted]	-	-	-
		Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	£20,348
2		Vehicle cream	[redacted]	[redacted]	-	-	-

No	Model scenario	Treatment	Total costs (£)	Total QALYs	Incremental Costs (£)	Incremental QALYs	ICER vs vehicle cream
	Utility data source: VNS	Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	£398,929
3	Model time horizon: 10 years	Vehicle cream	[redacted]	[redacted]	-	-	-
		Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	£5,687
4	Costs in the non-response state stop at: 5 years	Vehicle cream	[redacted]	[redacted]			
		Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	£39,272
5	Costs in the non-response state stop at: Lifetime	Vehicle cream	[redacted]	[redacted]	-	-	-
		Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	£3,894
6	Population: Overall	Vehicle cream	[redacted]	[redacted]	-	-	-
		Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	£19,179
7	Population: Fitzpatrick skin type IV-VI	Vehicle cream	[redacted]	[redacted]	-	-	-
		Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

Abbreviations: DP, depigmentation; F-VASI, facial vitiligo area scoring index; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year.

5.3. Model validation and face validity check

In B.3.11 of the CS, the company reported conducting validation exercises with “clinicians and health economists” throughout model conceptualisation, development and finalisation, though no evidence of such validation or what it comprised was provided in the CS. In response to priority EAG question B4, the company provided documentation from meetings with three anonymised “Clinical Expert”s and one anonymised “Health Economist”. Each expert was interviewed separately, in the presence of at least six attendees from the company and their consultancy. Two of the clinical experts were interviewed once, one was reinterviewed; these interviews each lasted 90-120 minutes and were conducted between January 2022 and June 2023. The health economist was interviewed three times between March 2022 and May 2023, in meetings lasting 60-120 minutes. The company appeared to have selectively used advice from these meetings. For example, the questions put to the clinical experts on response definitions and answers received across each clinical expert acknowledged the clinical relevance of VASI50 at 24 weeks. Yet, the company’s model categorised those achieving an “initial response” of VASI50-74 as non-reponders.

There was no evidence that the company validated model outcomes against published estimates from external studies or with clinical experts.

The company reported conducting a full quality control assessment following finalisation of the model, and provided as evidence the documentation produced by an internal quality control checklist exercise within Appendix N of the CS. Despite this being described as a full quality control, only some of the relevant checks in the document embedded in Appendix N appeared to have been conducted. The submitted model passed standard internal consistency and stress checks performed by the EAG as reported in section 4.2.2. However, sheet-by-sheet EAG review of the company model revealed clear issues with logic applied in the analysis, documented in section 4.2.5 and comprising part of Key Issue 3.

6. EXTERNAL ASSESSMENT GROUP'S ADDITIONAL ANALYSES

The EAG identified limitations within the company's base case and explored the impact of parameter values, and assumptions, which the EAG believed were more plausible.

This section is organised as follows: Section 6.1 details the impact of errors identified in the EAG's validation of the executable model. Section 6.2 details a series of scenario analyses exploring the robustness of the cost-effectiveness results to specific assumptions and additional uncertainties identified by the EAG. These analyses were conducted within the company corrected base-case analysis (presented in Section 6.1). The scenario analyses presented in Section 6.2 focus on exploring the following issues and uncertainties:

- Alignment of management costs with clinical practice (Key Issue 1 and Key Issue 5)
- Face validity of utility values (Key Issue 6)
- Dosing of ruxolitinib (Key Issue 4)
- Duration of costs application in 'no response' (Key Issue 5)
- Approach to handling missing data for clinical data (Key Issue 3)
- Retreatment with ruxolitinib (Key Issue 3)

As this list indicates, neither the EAG's scenario analyses nor the EAG-preferred analyses addressed every Key Issue identified throughout Sections 2, 3 and 4. Specifically, EAG amendments have not been able to address important elements of Key Issue 1 (The clinical and cost effectiveness of ruxolitinib as compared to established treatment options is unknown), Key Issue 2 (The clinical effectiveness evidence presented by the company was not representative of the target population and the population used in the company's economic evaluation), Key Issue 3 (Cost-effectiveness model's structural assumptions and use of clinical effectiveness data) and Key Issue 7 (Approach to adverse event assumptions in the cost-effectiveness model). As such, the EAG only presents *tentative* preferred analyses in this report.

In Section 6.3, the EAG's tentative preferred base-case results are based on a combination of the analyses presented in Section 6.2. Finally, Section 6.3 presents conclusions of the cost-effectiveness section of the EAG's report.

6.1. EAG corrections and adjustments to the company's base case model

The EAG corrections have addressed four errors in the company's model:

- The "initial period" of the model was intended to capture 6 cycles (i.e., 24 weeks), but instead captured 7 cycles (i.e., 28 weeks) – see Section 4.2.5. The EAG corrected this to limit the initial period to 6 cycles only. At factual accuracy check stage of the appraisal, the company highlighted that the same 'fix' should be applied for the long-term part of the model, and so the relevant cell values were reduced by 4 weeks (i.e., 1 cycle) accordingly.
- In response to clarification question B11, the company confirmed that since it populated its model, the BNF was updated which led to some costs no longer aligning with the stated sources – see Section 4.2.9. The EAG therefore amended these costs by updating cost sources to reflect current BNF costs using company's provided appendix, following cross-checking with the NHS eMIT for generic medicine costs (which were used in preference to the BNF if lower).
- The company's model included a calculation error where patients were mistakenly omitted from the numerator of a proportion calculation – see Section 4.2.7. The EAG corrected this.
- In Appendix M of the CS, the company provided variance-covariance matrices for utility regression analyses informing the company's health state utility assumptions. However, these were not integrated within the company's model to appropriately inform the probabilistic analysis. The EAG included the variance-covariance matrix, with implications for the probabilistic analysis.

The fixes are labelled as 'EAG_fix_1' to 'EAG_fix_4', and 'EAG_FAC_2', within the EAG-adapted model. When combined, these corrections collectively lead to a small reduction in the company's base-case deterministic ICER (£13,634 versus £13,031) and the mean probabilistic ICER (£14,676 versus £14,257), when the probabilistic analysis was based on 2,000 PSA iterations as in the company base case.

Table 26: EAG-corrected company base case results

	Discounted costs (£)	Discounted QALYs	Δ discounted costs (£)	Δ discounted QALYs	ICER (£/QALY)
EAG-corrected company deterministic base case					
Vehicle cream	[redacted]	[redacted]			

	Discounted costs (£)	Discounted QALYs	Δ discounted costs (£)	Δ discounted QALYs	ICER (£/QALY)
Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	13,031
EAG-corrected company mean probabilistic base case					
Vehicle cream	[redacted]	[redacted]			
Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	14,257

Abbreviations: ICER, incremental cost-effectiveness ratio; QALYs, quality adjusted life years.

6.2. Exploratory and sensitivity analyses undertaken by the EAG

Several exploratory analyses were undertaken to investigate the impact of alternative settings and assumptions on the cost-effectiveness results.

6.2.1. Alignment of management costs with clinical practice

As discussed in Sections 4.2.4 and 4.2.5, and central to Key Issue 1, the company's model may be considered to better reflect the potential use of ruxolitinib *following* second-line treatment options, rather than *between* first- and second-line options as the company proposed. This is because it does not provide a comparison to second-line treatment options, and instead presents a comparison to vehicle cream which is not a comparator relevant to this appraisal in this setting. In addition, based on clinical advice provided to the EAG, some of the cost assumptions informing the model were considered unlikely to reflect current NHS practice. For these reasons, the EAG considered the impact on results if resource use and cost assumptions model was edited to align with the use of ruxolitinib in an end-of-line setting (where vehicle cream could be considered as a proxy for 'no treatment', which may be considered a relevant comparator in this setting) and current NHS practice, which involved:

- Setting the cost of vehicle cream to £0 (since vehicle cream is not a treatment used in practice, and sun protection is already accounted for as part of concomitant therapy)
- Removing the cost of NB-UVB from the 'no-response' health state (given that this would be offered to patients in a second-line setting)
- Reducing the proportion of patients receiving psychological support to 15% (based on clinical advice provided to the EAG)

These changes were combined into one analysis, labelled 'EAG_1' in the model.

6.2.2. Face validity of utility values

The company's approach to estimating utility values, as described in Section 4.2.8, was both complex and subject to substantial limitations, which together contributed to Key Issue 6. As described within Section 4.2.8, the EAG had several key concerns with the company's base-case utility values. Two such concerns were investigated in EAG exploratory analyses:

- As illustrated by EAG clarification question B15, the expected utility values assumed for "Maintenance period", "Stable" and "Stable retreated" health states were higher than the age-equivalent general population utility value from a source commonly cited in NICE appraisals[37]. Notably, this same source was used in the company's own model to adjust utility for the effect of ageing over the model's time horizon.

- The company's central approach to utility value estimation produced a utility value of 0.890 for patients achieving F-VASI50-74 at 24 weeks, as documented in Table 21. Yet, in the company's model, patients achieving F-VASI50-74 are categorised as "Non-responders" and assigned a utility value of 0.797.

To address these issues, the EAG conducted the following two exploratory analyses:

1. capped all utility values using general population estimates and limited the reduction in utility from baseline to no response to 5% (i.e., the utility for the 'no response' health state was assumed to be 95% of the utility value for baseline).
2. capped all utility values using general population estimates and applied a weighted average of the utility value estimated for patients with F-VASI50-74 and 'no response' to the "non-response" state, using the proportion of patients in the ruxolitinib arm in each category measured at 24 weeks. Please note: at the factual accuracy check stage of the appraisal, the company highlighted that that EAG's analysis included a numerical error in determining the relevant weights, which was subsequently corrected.

These changes were included within the model, labelled 'EAG_2' and 'EAG_FAC_1' in the model, alongside a multi-way sensitivity analysis to consider alternative combinations of utility values.

6.2.3. Ruxolitinib dosing

The average dose of ruxolitinib used by people with vitiligo in practice was challenging to estimate for a variety of reasons, as discussed in Section 4.2.4. These included (but may not be limited to): differences in extent of depigmentation, body surface area, patient preference for treating specific regions, adherence to treatment, and interpretation of dosing instructions. In the company's base-case analysis, the median dose across both treatment arms across the double-blind period of the TRuE-V studies was assumed to represent expected ruxolitinib use. This gave an average daily dose of 4.03g, which was equivalent to 1.13 tubes per 28 days. As described in Section 4.2.4 and comprising Key Issue 4, the EAG was concerned that the company's dosing assumptions may substantially underestimate ruxolitinib use, in a manner that biased cost-effectiveness results notably.

The EAG would generally prefer to use the mean dose for patients only receiving ruxolitinib (i.e., excluding dosing data for the vehicle cream arm) across both the TRuE-V studies. The mean dose for this sample, as reported by the company in response to clarification question B10, was 7.61g daily; equivalent to 2.13 tubes per 28 days. Given that the SmPC for ruxolitinib advised that no more than two tubes should be used per patient per month, an alternative estimate of the average (and maximum) daily dose would be 6.57g, which was equivalent to exactly two tubes per 30.4375 days. Consequently, in two alternative analyses, the EAG applied either a daily dose of 6.57g or the mean daily dose of 7.61g, acknowledging that the latter exceeds the advised upper limit of two tubes per month, but nevertheless represented the observed average use of ruxolitinib in the TRuE-V studies. Furthermore, the EAG noted that the average dose of ruxolitinib appeared to increase in the TRuE-V LTE study (see company's response to clarification question B10).

Finally, the EAG also considered a scenario analysis in which two tubes are provided to patients every 28 days, in keeping with 28-day prescribing patterns, the model cycle length, and the guidance given to patients as part of the TRuE-V studies (i.e., 60g per 7 days). This scenario was introduced following a correction made to the interpretation of "two tubes per month" at the factual accuracy check stage of the appraisal.

Alternative dosing assumptions were incorporated into the model, labelled as 'EAG_3' and 'EAG_FAC_3'.

6.2.4. Duration of costs application in 'non-response'

The company's model assumed that all costs in the 'non-response' state would cease from 10 years following model entry, without clear and complete rationale. It was the EAG's view that 'time to no further costs' would be linked with residence in and time since arrival to the 'non response' state, as opposed to how much time has elapsed since model entry. Therefore, while the EAG acknowledged that some people may, in time, become disengaged with secondary care, the company did not substantiate its assumption with respect to the application of 10 years from model entry.

In a scenario analysis, the EAG removed the 10-year cap on accrual of costs in the 'no response' health state. This is labelled as 'EAG_4' in the model. In an alternative analysis, the EAG also considered applying 10 years' worth of costs for the proportion of patients that were no longer in a response or treated health state compared with the previous model cycle. This is labelled as 'EAG_sc_2' in the model.

The EAG presents these pragmatic exploratory analyses in lieu of a better alternative, as it would take a substantial amount of modelling work to specify tunnel states to track time since entry to the 'no response' health state, which would not be feasible within the timeframe of the EAG appraisal.

6.2.5. Approach to handling missing data for no regain of response

As noted in Section 4.2.7 and comprising part of Key Issue 3, the company's approach to handling missing data for the probability of not regaining response was erroneous. In addition, the EAG considered a more accurate estimate of this probability to be based on the assumption that missing data were likely to reflect non-response. Therefore, in a scenario conducted by the EAG (labelled 'EAG_5'), an alternative probability of not regaining response was used in place of the company's preferred 'average' approach.

6.2.6. Retreatment with ruxolitinib

The company's model structure assumed that patients who achieved F-VASI90 and became "Stable" but subsequently lost this level of repigmentation move to the 'Retreatment' health state (see Figure 5). At the clarification stage, the EAG asked the company to update its model to allow for an analysis in which retreatment for such patients was not assumed to be certain (clarification question B7). In response, the company included the ability for the proportion of

patients assumed to be re-treated to be varied between 0% and 100%, including a scenario using a value of 68% (though no citation was provided for this value).

Clinical advice to the EAG suggested that all patients would be offered re-treatment if they previously were deemed to respond sufficiently well to treatment following an initial course. However, as highlighted in Section 4.2.5 of this report, the expected continuation criteria used in NHS clinical practice may differ from the criteria used to determine transitions in the model structure. In addition, it remained unclear how many courses of treatment with ruxolitinib patients may undergo over their lifetime, and it was unclear if or how the effect of ruxolitinib may change if it was used in successive courses (especially accounting for patients starting treatment with likely differing extent of de-pigmentation compared with baseline). Consequently, the EAG noted that it was difficult to justify a preferred base-case setting for this parameter, within the constraints of the company's model.

In the absence of any clear rationale to deviate from the company's base-case setting of 100% retreatment (which in principle was supported by clinical advice to the EAG, notwithstanding the differences in continuation criteria as previously noted), the EAG adhered to the company's base-case assumption within its tentatively preferred base-case analysis. In sensitivity analysis, the EAG disabled re-treatment entirely to ascertain its impact on the cost-effectiveness results. This is labelled as 'EAG_sc_1' in the model.

6.2.7. Impact on the ICER of additional clinical and economic analyses undertaken by the EAG

The EAG made the changes described in Sections 6.2.1 to 6.2.6 (labelled as 'EAG_1' to 'EAG_5', and 'EAG_sc_1' to 'EAG_sc_2' in the model). The results of these exploratory analyses (where each change has been made individually) are provided in Table 27.

Table 27: EAG's exploratory analyses with EAG fixes applied

Scenario	Section	Δ costs (£)	Δ QALYs	ICER (£/QALY)	+/- company base case
EAG-corrected deterministic company base case	6.1	[redacted]	[redacted]	13,031	-603
Alignment of management costs with clinical practice – removal of vehicle cream and NB-UVB costs, removal of dermatology visits for patients in 'no response' and	6.2.1	[redacted]	[redacted]	99,237	+85,603

Scenario	Section	Δ costs (£)	Δ QALYs	ICER (£/QALY)	+/- company base case
proportion of patients receiving psychological support set to 15% for all health states (EAG_1)					
Face validity of utility values – maximum utilities set to those of the general population and 'no response' utilities set to a weighted average of baseline and F-VASI50-74 values (EAG_2 and EAG_FAC_1)	6.2.2	[redacted]	[redacted]	21,640	+8,005
Face validity of utility values – maximum utilities set to those of the general population and 'no response' utilities set to an arbitrary reduction of 5% from baseline utility values (EAG_2)	6.2.2	[redacted]	[redacted]	25,822	+12,188
Ruxolitinib dosing – assume pooled mean dose for only the ruxolitinib arms of TRuE-V studies (EAG_3)	6.2.3	[redacted]	[redacted]	96,046	+82,412
Ruxolitinib dosing – assume a maximum recommended daily dose when two tubes per month (EAG_3)	6.2.3	[redacted]	[redacted]	71,894	+58,260
Ruxolitinib dosing – assume a maximum recommended daily dose when two tubes per 28 days (EAG_FAC_3)	6.2.3	[redacted]	[redacted]	85,146	+71,512
Duration of costs application in 'non-response' – set to lifetime (EAG_4)	6.2.4	[redacted]	[redacted]	3,567	-10,067
Duration of costs application in 'non-response' – alternative application based on lump sum (EAG_sc_2)	6.2.4	[redacted]	[redacted]	78,252	+64,618
Approach to handling missing data for no regain of response – assume non-response (EAG_5)	6.2.5	[redacted]	[redacted]	13,580	-54
Retreatment with ruxolitinib set to 0% (EAG_sc_1)	6.2.6	[redacted]	[redacted]	31,354	+17,720

Abbreviations: EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

6.3. EAG's preferred assumptions

As described in Section 6.2, the EAG presents two preferred analyses that differ only in their dosing assumptions, and considered all analyses "tentative", given the outstanding and unexplored uncertainty around Key Issues 1, 2, 3 and 7, collectively.

The following deviations from the EAG-corrected company base case applied to both EAG-preferred tentative base cases:

- Removal of vehicle cream and NB-UVB costs, removal of dermatology visits for patients in 'no response' and proportion of patients receiving psychological support set to 15% for all health states to align with clinical practice (Section 6.2.1; Key Issue 1 and Key Issue 5)
- Maximum utility values set to those of the general population, and 'no response' values set to a weighted average of baseline and F-VASI50-74 utility values (Section 6.2.2; Key Issue 6)
- Duration of costs applied for both drug acquisition and disease management set to lifetime (Section 6.2.4; Key Issue 3 and Key Issue 5)
- Missing data treated as non-response data in calculation of the probability of retreatment (given F-VASI<75 at week 52 leads to F-VASI<90 at week 104; Section 6.2.5; Key Issue 3)

EAG preferred, tentative base cases 1 and 2 differ with respect to expected ruxolitinib use assumptions only. Tentative Base Case 1 results, shown across Table 28 and Table 29, assume that ruxolitinib use in practice will reflect the pooled mean dose for the ruxolitinib arms of TRuE-V studies (7.61g). Tentative Base Case 2 results, shown across Table 30 and Table 31, assume ruxolitinib use in practice will be limited to the SmPC recommendation of no more than two 100g tubes per patient per month (6.57g).

EAG adjustments collectively reduce the expected incremental QALY gain associated with ruxolitinib while increasing its expected incremental cost, leading to EAG-preferred tentative ICERs that were far in excess of the relevant NICE decision-making threshold range, as results across Table 28, Table 29, Table 30 and Table 31 show. Figure 9 and Figure 10 serve to illustrate this point; in PSAs, all PSA iterations were above the £20,000 per QALY willingness to pay threshold. Mean probabilistic ICERs were higher than deterministic ICERs, owing to the skewed distribution of PSA iterations visible in Figure 9 and Figure 10. This trend was not present in the company's probabilistic analysis, and as such was likely attributable to EAG

correction #4, which incorporated variance-covariance matrices for utility regression analyses informing the company's health state utility assumptions into the cost-effectiveness model.

Table 28: From EAG-corrected company base case results to EAG-preferred tentative base case 1 results (all deterministic)

Preferred assumption	Section in EAG report	Cumulative ICER £/QALY
EAG-corrected deterministic company base case	6.1	*13,031*
Alignment of management costs with clinical practice – removal of vehicle cream and NB-UVB costs, removal of dermatology visits for patients in 'no response' and proportion of patients receiving psychological support set to 15% for all health states	6.2.1	*99,237*
Face validity of utility values – maximum utilities set to those of the general population and 'no response' utilities set to a weighted average of baseline and F-VASI50-74 values	6.2.2	*164,794*
Ruxolitinib dosing – assume pooled mean dose for only the ruxolitinib arms of TRuE-V studies	6.2.3	*302,651*
Duration of costs application in 'non-response' – set to lifetime	6.2.4	*301,699*
Approach to handling missing data for no regain of response – assume non-response	6.2.5	*303,189*

Abbreviations: EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

Table 29: Summary EAG-preferred tentative base case 1 results

	Discounted costs (£)	Discounted QALYs	Δ discounted costs (£)	Δ discounted QALYs	ICER (£/QALY)
Deterministic					
Vehicle cream	[redacted]	[redacted]			
Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	303,189
Probabilistic					
Vehicle cream	[redacted]	[redacted]			
Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	329,105

Abbreviations: EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALYs, quality adjusted life years

Figure 9: PSA scatterplot, EAG-preferred tentative base case 1

Abbreviations: EAG, External Assessment Group; PSA, probabilistic sensitivity analysis; QALYs, quality adjusted life years; WTP, willingness-to-pay

Table 30: From EAG-corrected company base case results to EAG-preferred tentative base case 2 results (all deterministic)

Preferred assumption	Section in EAG report	Cumulative ICER £/QALY
EAG-corrected deterministic company base-case	6.1	*13,031*
Alignment of management costs with clinical practice – removal of vehicle cream and NB-UVB costs, removal of dermatology visits for patients in 'no response' and proportion of patients receiving psychological support set to 15% for all health states	6.2.1	*99,237*
Face validity of utility values – maximum utilities set to those of the general population and 'no response' utilities set to a weighted average of baseline and F-VASI50-74 values	6.2.2	*164,794*
Ruxolitinib dosing – assume maximum dose according to SmPC recommendation for 2 x 100g tubes per patient per month	6.2.3	*262,543*
Duration of costs application in 'non-response' – set to lifetime	6.2.4	*261,592*
Approach to handling missing data for no regain of response – assume non-response	6.2.5	*262,880*

Abbreviations: EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

Table 31: Summary EAG-preferred tentative base case 2 results

	Discounted costs (£)	Discounted QALYs	Δ discounted costs (£)	Δ discounted QALYs	ICER (£/QALY)
Deterministic					
Vehicle cream	[redacted]	[redacted]			
Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	262,880
Probabilistic					
Vehicle cream	[redacted]	[redacted]			
Ruxolitinib cream	[redacted]	[redacted]	[redacted]	[redacted]	283,278

Abbreviations: EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALYs, quality adjusted life years

Figure 10: PSA scatterplot, EAG-preferred tentative base case 2

Abbreviations: PSA, probabilistic sensitivity analysis; QALYs, quality adjusted life years; WTP, willingness-to-pay

6.3.1. Scenario analyses around the EAG's preferred assumptions

Table 32: Individual impact of each scenario upon the EAG's and Table 36 present univariate scenario analyses around the EAG-preferred tentative base case results. The tables are a reflection of the exploratory analyses around the EAG-corrected company base case results

shown in Table 27, except around EAG-preferred results. They serve to illustrate the isolated importance of relaxing each of the EAG's proposed changes.

Table 32: Individual impact of each scenario upon the EAG's deterministic Base Case 1 ICER

Scenario	Section	Δ costs (£)	Δ QALYs	ICER (£/QALY)	+/- EAG base case
Base Case 1	-	[redacted]	[redacted]	303,189	-
Alignment of management costs with clinical practice – reverting to company assumptions	6.2.1	[redacted]	[redacted]	145,374	-157,815
Face validity of utility values – reverting to company assumptions	6.2.2	[redacted]	[redacted]	182,586	-120,603
Face validity of utility values – maximum utilities set to those of the general population and 'no response' utilities set to an arbitrary reduction of 5% from baseline utility values	6.2.2	[redacted]	[redacted]	361,765	+58,575
Ruxolitinib dosing – reverting to company assumptions	6.2.3	[redacted]	[redacted]	164,638	-138,551
Ruxolitinib dosing – assume a maximum recommended daily dose when two tubes per month	6.2.3	[redacted]	[redacted]	262,880	-40,310
Ruxolitinib dosing – assume a maximum recommended daily dose when two tubes per 28 days	6.2.3	[redacted]	[redacted]	284,996	-18,193
Duration of costs application in 'non-response' – reverting to company assumptions	6.2.4	[redacted]	[redacted]	304,141	+952
Duration of costs application in 'non-response' – alternative application based on lump sum	6.2.4	[redacted]	[redacted]	310,694	+7,505
Approach to handling missing data for no regain of response – reverting to company assumptions	6.2.5	[redacted]	[redacted]	301,699	-1,490
Retreatment with ruxolitinib set to 0%	6.2.6	[redacted]	[redacted]	350,808	+47,619

Abbreviations: EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

Table 33: Individual impact of each scenario upon the EAG's deterministic Base Case 2 ICER

Scenario	Section	Δ costs (£)	Δ QALYs	ICER (£/QALY)	+/- EAG base case
Base Case 2	-	[redacted]	[redacted]	262,880	-
Alignment of management costs with clinical practice – reverting to company assumptions	6.2.1	[redacted]	[redacted]	105,064	-157,815
Face validity of utility values – reverting to company assumptions	6.2.2	[redacted]	[redacted]	158,311	-104,569
Face validity of utility values – maximum utilities set to those of the general population and 'no response' utilities set to an arbitrary reduction of 5% from baseline utility values	6.2.2	[redacted]	[redacted]	313,667	+50,788
Ruxolitinib dosing – assume pooled mean dose for only the ruxolitinib arms of TRuE-V studies	6.2.3	[redacted]	[redacted]	303,189	+40,310
Ruxolitinib dosing – reverting to company assumptions	6.2.3	[redacted]	[redacted]	164,638	-98,241
Ruxolitinib dosing – assume a maximum recommended daily dose when two tubes per 28 days	6.2.3	[redacted]	[redacted]	284,996	+22,117
Duration of costs application in 'non-response' – reverting to company assumptions	6.2.4	[redacted]	[redacted]	263,832	+952
Duration of costs application in 'non-response' – alternative application based on lump sum	6.2.4	[redacted]	[redacted]	270,384	+7,505
Approach to handling missing data for no regain of response – reverting to company assumptions	6.2.5	[redacted]	[redacted]	261,592	-1,288
Retreatment with ruxolitinib set to 0%	6.2.6	[redacted]	[redacted]	304,039	+41,159

Abbreviations: EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

In addition to the results presented in the tables above, multi-way sensitivity analyses around the tentative EAG-preferred base cases were performed for utility values, to establish the impact of jointly varying the utility values assumed for F-VASI90, F-VASI75-89, and 'non-response' health states. The details of and results from these analyses are reported in Appendix B.

6.4. Conclusions of the cost-effectiveness section

The EAG was not satisfied that the cost-effectiveness evidence submitted by the company addressed the decision problem at hand. The company's analysis used clinical effectiveness data from the previously treated subgroups of its pivotal registrational trials, and so addressed a subgroup of the licensed population and the final scope. This in itself was surmountable, if for example, the revised target population was considered identifiable and definable in guidance, although the EAG also noted the lack of comprehensive information about this subgroup presented in the CS (Key Issue 2). However, the cost-effectiveness comparison the company presented, to vehicle cream, was not appropriate for the proposed population. As such, the EAG did not find the company's model useful for addressing the decision problem the company proposed. The EAG considered that a comparison to vehicle cream, as a proxy for no active treatment, may only be appropriate for an end-of-line setting.

Furthermore, the EAG was not satisfied that the company's cost-effectiveness results provided an unbiased estimate of the likely cost-effectiveness of ruxolitinib. Most notably, the EAG identified: issues with the company's model logic and use of clinical effectiveness data (Key Issue 3); evidence that the company's dosing assumptions were underestimating expected ruxolitinib costs (Key Issue 4); indications that the company's preferred "non-response" state resource use assumptions overestimated healthcare costs (Key Issue 5); plausibility and internal consistency issues with the company's preferred health state utility estimates (Key Issue 6); an approach to capture adverse event consequences that underestimated cost and did not consider patient utility implications (Key Issue 7).

Within the timeframe of the EAG appraisal, the EAG was able to resolve some but not all of the Key issues identified. The EAG therefore present only "tentative" preferred results. These results, albeit tentative, did not suggest that the expected health benefits of ruxolitinib were sufficient to justify its expected incremental costs, given decision-making thresholds. This finding contrasted starkly with the company-preferred results.

Overall, the cost-effectiveness of ruxolitinib was highly uncertain. Substantial uncertainty could be resolved if the company addressed the outstanding Key Issues documented in this report.

References

1. Electronic Medicines Compendium. Opzelura 15 mg/g cream: summary of product characteristics 2023. Available from: https://www.medicines.org.uk/emc/product/14903/smpc.
2. Oiso N, Suzuki T, Fukai K, Katayama I, Kawada A. Nonsegmental vitiligo and autoimmune mechanism. Dermatol Res Pract. 2011;2011:518090.
3. Thomas KS, Batchelor JM, Akram P, Chalmers JR, Haines RH, Meakin GD, et al. Randomized controlled trial of topical corticosteroid and home-based narrowband ultraviolet B for active and limited vitiligo: results of the HI-Light Vitiligo Trial. Br J Dermatol. 2021;184(5):828-39.
4. Incyte. INCB - 88888-022 Clinical Study Report: REVEAL-UK: “Real world Evidence in Vitiligo Employing dAtabase anaLyses in UK” [Data on file]. 2022.
5. Incyte. Summary of clinical validation discussions. 2023.
6. van Geel N, Speeckaert M, Brochez L, Lambert J, Speeckaert R. Clinical profile of generalized vitiligo patients with associated autoimmune/autoinflammatory diseases. J Eur Acad Dermatol Venereol. 2014;28(6):741-6.
7. TRuE-V1 I. Topical Ruxolitinib Evaluation in Vitiligo Study 1 (TRuE-V1): A Phase 3, Double-Blind, Randomized, Vehicle-Controlled, Efficacy and Safety Study of Ruxolitinib Cream Followed by an Extension Period in Participants With Vitiligo. 2021.
8. TRuE-V2 I. A Phase 3, Double-Blind, Randomized, Vehicle-Controlled, Efficacy and Safety Study of Ruxolitinib Cream Followed by an Extension Period in Participants With Vitiligo. 2021.
9. Incyte. INCB 18424-308 Clinical Study Report: A Double-Blind, Vehicle-Controlled, Randomized Withdrawal and Treatment Extension Study to Assess the Long-Term Efficacy and Safety of Ruxolitinib Cream in Participants With Vitiligo (TRuE-V LTE) [Data on file]. 2023.
10. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. bmj. 2016;355.
11. Incyte. Efficacy and Safety of Clinical interventions in Vitiligo: A Systematic Literature Review [Data on file]. 2023.
12. Incyte. INCB 18424-211 clinical study report [Data on file]. 2019.
13. Incyte. A Study to Evaluate the Mechanism of Action of Ruxolitinib Cream in Subjects With Vitiligo (TRuE-V MOA) 2021. Available from: https://classic.clinicaltrials.gov/ct2/show/record/NCT04896385?view=record.
14. Incyte. NCT02809976 - Topical Ruxolitinib for the Treatment of Vitiligo [Completed]. Phase 2. 2016. Available from: https://classic.clinicaltrials.gov/ct2/show/NCT02809976.
15. Incyte. NCT05750823 - a study to assess the safety and efficacy of ruxolitinib cream in participants with genital vitiligo. 2023. Available from: https://classic.clinicaltrials.gov/ct2/show/NCT05750823.
16. Incyte. NCT05247489 A Study to Evaluate the Safety and Efficacy of Ruxolitinib Cream With Phototherapy in Participants With Vitiligo. 2022. Available from: https://classic.clinicaltrials.gov/ct2/show/NCT05247489.
17. Centre Hospitalier Universitaire de Nice. NCT05872477 Promoting Repigmentation After Epidermal Cell Suspension Grafting and preVENTing the Loss of Melanocytes Using Topical Ruxolitinib for Vitiligo in Resistant Areas (PREVENT). 2023. Available from: https://classic.clinicaltrials.gov/ct2/show/NCT05872477.
18. Incyte. INCB 18424-306 Clinical Study Report: Topical Ruxolitinib Evaluation in Vitiligo Study 1 (TRuE-V1) [Data on file]. 2022.
19. Incyte. INCB 18424-307 Clinical Study Report: Topical Ruxolitinib Evaluation in Vitiligo Study 2 (TRuE-V2) [Data on file]. 2022.

20. Office for National Statistics. Ethnic group, England and Wales: Census 2021 2022. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/bulletins/ethnicgroupenglandandwales/census2021.
21. Eleftheriadou V, Atkar R, Batchelor J, McDonald B, Novakovic L, Patel JV, et al. British Association of Dermatologists guidelines for the management of people with vitiligo 2021. Br J Dermatol. 2022;186(1):18-29.
22. Kitchen H, Wyrwich KW, Carmichael C, Deal LS, Lukic T, Al-Zubeidi T, et al. Meaningful Changes in What Matters to Individuals with Vitiligo: Content Validity and Meaningful Change Thresholds of the Vitiligo Area Scoring Index (VASI). Dermatol Ther (Heidelb). 2022;12(7):1623-37.
23. Lilly E, Lu PD, Borovicka JH, Victorson D, Kwasny MJ, West DP, et al. Development and validation of a vitiligo-specific quality-of-life instrument (VitiQoL). J Am Acad Dermatol. 2013;69(1):e11-8.
24. Centre for Reviews and Dissemination. CRD's guidance for undertaking reviews in healthcare: York Publ. Services; 2009.
25. Incyte. Psychometric Analysis Report. Phase 3, Double-Blind, Randomized, 8-Week, Vehicle-Controlled Efficacy and Safety Studies of Ruxolitinib Cream Followed by a Long-Term Safety Extension Period in Participants with Vitiligo (TRuE-V1 and TRuE-V2). Incyte report number ED-21.01.1 [Data on file]. 2021.
26. EMA. Opzelura assessment report (EMA/135534/2023). 2023.
27. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatrica Scandinavica. 1983;67(6):361-70.
28. Lemay KR, Tulloch HE, Pipe AL, Reed JL. Establishing the Minimal Clinically Important Difference for the Hospital Anxiety and Depression Scale in Patients With Cardiovascular Disease. Journal of Cardiopulmonary Rehabilitation and Prevention. 2019;39(6):E6-E11.
29. Lin JQ, Li SQ, Li S, Kiamanesh EF, Aasi SZ, Kwong BY, et al. A 10-year retrospective cohort study of ruxolitinib and association with nonmelanoma skin cancer in patients with polycythemia vera and myelofibrosis. J Am Acad Dermatol. 2022;86(2):339-44.
30. Begum R, Crott R, Martina R, Loizidou EM, Khan I. Estimating health related quality of life effects in vitiligo. Mapping EQ-5D-5 L utilities from vitiligo specific scales: VNS, VitiQoL and re-pigmentation measures using data from the HI-Light trial. Health and Quality of Life Outcomes. 2023;21(1):85.
31. Eleftheriadou V, Thomas, K., Ravenscroft, J., Whitton, M., Batchelor, J. and Williams, H. Feasibility, double-blind, randomised, placebo-controlled, multi-centre trial of hand-held nb-uvb phototherapy for the treatment of vitiligo at home (hi-light trial: Home intervention of light therapy). Trials. 2014;15(1):1-12.
32. Rosmarin D, Passeron T, Pandya AG, Grimes P, Harris JE, Desai SR, et al. Two Phase 3, Randomized, Controlled Trials of Ruxolitinib Cream for Vitiligo. N Engl J Med. 2022;387(16):1445-55.
33. NICE. Guide to the methods of technology appraisal. Process and methods [PMG9]. 2013.
34. Hernández Alava M, Pudney S, Wailoo A. Estimating the Relationship Between EQ-5D-5L and EQ-5D-3L: Results from a UK Population Study. PharmacoEconomics. 2023;41(2):199-207.
35. NICE. NICE health technology evaluations: the manual Process and methods. 2022.
36. Incyte. Technical Report for Statistical Analysis and Utility Modelling of Data from two Randomized Controlled Trials for the treatment of Vitiligo to support HTA of Ruxolitinib Cream. Data on file 2023.
37. Ara RB, J. . Populating an economic model with health state utility values: moving towards better practice.: The University of Sheffield,; 2010.

38. Hamzavi IH, Bibeau K, Grimes P, Harris JE, van Geel N, Parsad D, et al. Exploring the natural and treatment history of vitiligo: patient and healthcare professional perceptions from the global VALIANT study. Br J Dermatol. 2023.
39. Sach TH, Thomas KS, Batchelor JM, Perways A, Chalmers JR, Haines RH, et al. An economic evaluation of the randomized controlled trial of topical corticosteroid and home-based narrowband ultraviolet B for active and limited vitiligo (the HI-Light Vitiligo Trial)*. British Journal of Dermatology. 2021;184(5):840-8.
40. NHS England. NHS Talking Therapies, for anxiety and depression. Available from: https://www.england.nhs.uk/mental-health/adults/nhs-talking-therapies/.
41. European Medicines Agency. NOTE FOR GUIDANCE ON STRUCTURE AND CONTENT OF CLINICAL STUDY REPORTS (CPMP/ICH/137/95). London; 1996.
42. Ezzedine K, Eleftheriadou V, Jones H, Bibeau K, Kuo FI, Sturm D, et al. Psychosocial Effects of Vitiligo: A Systematic Literature Review. Am J Clin Dermatol. 2021;22(6):757-74.
43. Nanja van Geel M, John E. Harris, MD, PhD, Iltefat H. Hamzavi, MD, Kristen Bibeau, PhD, MSPH, Jessy Gao, MA, Haobo Ren, PhD, Khaled Ezzedine, MD, PhD. Exploring the Natural and Treatment History of Vitiligo in Europe: Findings From the Global VALIANT StudyP1401Presented. 31st European Academy of Dermatology and Venereology (EADV) Congress. 2022.
44. Merhi S, Salameh P, Abboud M, Seneschal J, Eleftheriadou V, Pane I, et al. Facial involvement is reflective of patients' global perception of vitiligo extent. Br J Dermatol. 2023;189(2):188-94.

Appendix A: EAG indirect treatment comparison feasibility analysis

Table 34. EAG indirect treatment comparison feasibility analysis

	TRuE-V1[18]/TRuE-V2[19,32]	Thomas 2021 (HI-Light Vitiligo Trial){Thomas, 2021 #111	Eleftheriadou 2014[31]
	Inclusion / exclusion criteria		
Inclusion criteria	- Patients aged ≥ 12 years with a clinical diagnosis of nonsegmental vitiligo - Depigmented areas including ≥ 0.5% BSA on the face - ≥ 0.5 F-VASI - ≥ 3% BSA on nonfacial areas - ≥ 3 T-VASI - Total body vitiligo area (facial and nonfacial) was not to exceed 10% BSA	- Patients 5 years of age or over with a diagnosis of non-segmental vitiligo confirmed by a dermatologist. - Vitiligo limited to approximately 10% or less of body surface area, with at least one patch that is reported by the participant to have been active in the last 12 months. - No other active therapy for vitiligo (or willing to stop current treatment – no washout period required). - Able to administer the intervention safely at home	- Patients 5 years of age or over with a diagnosis of non-segmental vitiligo confirmed by a dermatologist. - Vitiligo affecting less than 25% or less of body surface area - No therapy for vitiligo in the previous two weeks and no other concurrent vitiligo treatments during the trial were allowed.
Exclusion criteria	- Other types of vitiligo - Patients who had no pigmented hair within any of the vitiligo areas on the face. - Patients who had used depigmentation treatments (e.g., monobenzone) - Any other skin disease that would interfere with the study medication application or study assessments - Any serious illness or medical, physical, or psychiatric condition(s) that would interfere with full participation in the study	- Other types of vitiligo (e.g. segmental or universal vitiligo). - History of skin cancer - History of radiotherapy use - Photosensitivity - Current use of immunosuppressive drugs	- Segmental or universal Vitiligo - Previous history of skin cancer - Recent/concurrent radiotherapy - Photosensitivity - Use of immunosuppressive or photosensitive drugs
	Actual distribution of demographics/disease characteristics between sources		

	TRuE-V1[18]/TRuE-V2[19,32]		**Thomas 2021 (HI-Light Vitiligo Trial){Thomas, 2021 #111**			**Eleftheriadou 2014[31]**	
Treatment arm	Ruxolitinib cream	Placebo (vehicle cream)	Placebo device with TCS	NB-UVB with placebo cream	NB-UVB with TCS	NB-UVB	Placebo device
Number of patients	450	224	173	169	175	19	10
Mean (SD) Age in years	39.5 (15.38)	39.7(14.5)	38.6 (20.0)	36.9 (18.9)	37.0 (19.1)	27.6 (18.6)	39.4 (13.5)
Min, max	12, 79	12, 79	NR	NR	NR	5, 71	13, 51
Sex, n (%)							
Male	202 (44.9)	114 (50.9)	75 (43.3)	88 (52.1)	105 (60.0)	10 (52.6)	5 (50.0)
Female	248 (55.1)	110 (49.1)	98 (56.6)	81 (48.0)	70 (40.0)	9 (47.4)	5 (50.0)
Fitzpatrick skin type, n (%)							
I	12 (2.7)	4 (1.8)	2 (1.2)	2 (1.2)	5 (2.9)	NR	NR
II	131 (29.1)	72 (32.1)	31 (17.9)	32 (18.9)	29 (10.9)	NR	NR
III	179 (39.8)	88 (39.3)	70 (40.4)	66 (39.1)	59 (19.4)	NR	NR
IV	89 (19.8)	40 (17.9)	29 (16.8)	34 (20.1)	33 (18.9)	NR	NR
V	28 (6.2)	17 (7.6)	35 (20.2)	25 (14.8)	44 (25.1)	NR	NR
VI	11 (2.4)	3 (1.3)	6 (3.5)	10 (5.9)	10 (5.7)	NR	NR
Race, n (%)							
White	363 (80.7)	189 (84.4)	112 (64.7)	114 (67.5)	104 (59.4)	12 (63.2)	8 (80.0)
Black/African American	23 (5.1)	9 (4.0)	5 (2.9)	3 (1.8)	7 (4.0)	2 (10.5)	0
Asian	17 (3.8)	11 (4.9)	36 (20.8)	39 (23.1)	49 (28.0)	3 (15.8)	2 (20.0)
American Indian/Alaska Native	2 (0.4)	0	NR	NR	NR	NR	NR
Native Hawaiian/Pacific Islander	2 (0.4)	0	NR	NR	NR	NR	NR
Mixed race	NR	NR	9 (5.2)	6 (3.6)	6 (3.4)	1 (5.3)	0

	TRuE-V1[18]/TRuE-V2[19,32]		Thomas 2021 (HI-Light Vitiligo Trial){Thomas, 2021 #111			Eleftheriadou 2014[31]	
Not reported	19 (4.2)	6 (2.7)	1 (0.06)	0	1 (0.06)	0	0
Other	24 (5.3)	9 (4.0)	10 (5.8)	7 (4.1)	9 (5.1)	1 (5.3)	0
Years since initial diagnosis of vitiligo							
Mean (SD)	14.9 (11.9)	14.6 (11.0)	NR	NR	NR	11.4 (10.1)	14.0 (8.5)
Median (IQR)	11.8	12.1	7 (3-16)	5 (3–11)	7 (4–15)	NR	NR
Disease status, n (%)							
Stable	331 (73.6)	168 (75.0)	NR[a]	NR[a]	NR[a]	5 (26.3) – ‘stable’ or ‘repigmenting’	5 (50.0) – ‘stable’ or ‘repigmenting’
Progressive	119 (26.4)	56 (25.0)	NR[a]	NR[a]	NR[a]	14 (76.7) – “spreading”	5(50.0) – “spreading”
T-BSA involvement							
Mean (SD)	7.36 (2.0)	7.46 (2.0)	NR	NR	NR	9.8 (6.0)	6.9 (6.2)
Prior therapy received, n (%)							
Topical corticosteroids	133 (29.6)	56 (25.0)	80 (46.2)	75 (44.4)	80 (45.6)	NR	NR
Topical calcineurin inhibitor	146 (32.4)	68 (30.4)	51 (29.5)	39 (23.1)	56 (32.0)	NR	NR
NB-UVB	138 (30.7)	77 (34.4)	28 (16.2) received “light therapy”	26 (15.4) received “light therapy”	37 (21.1) received “light therapy”	NR	NR
Outcomes							
Patient reported vitiligo scales	Patient reported VNS at 24 weeks and 52 weeks (5 point scale)		Patient reported VNS at 9 months.			Global improvement in vitiligo: 5-point Likert scale	
Repigmentation	T-VASI75[b] / F-VASI75[b] at 24 weeks and 52 weeks		≥ 75% repigmentation at 9 months on the “target patch”			≥ 75% repigmentation at 16 weeks in up to 3 target lesions per patient	

[a] The study had an inclusion criteria was having a vitiligo patch that was reported as active (new or changed). Thus, it would appear all participants has progressive vitiligo.

[b] T-VASI75/F-VASI75: achieving at least 75% improvement from baseline.

Abbreviations**:** BSA, body surface area; F-VASI, face vitiligo area scoring index; SD, standard deviation; T-BSA, total body surface area; TCI, topical calcineurin

Appendix B: Multi-way sensitivity analysis of utility values

As noted in Section 6.2.2, there is substantial uncertainty concerning the estimation of utility values to populate the company's model. Given the differences between the company's preferred utility values and the EAG's preferred utility values (see Section 6.2.2), a further multi-way analysis was conducted. In the multi-way sensitivity analysis:

- The company's preferred utility values were considered upper limits, whereas the EAG's preferred values were considered lower limits.
- Utility for 'no response' was varied between 0.797 and 0.822, in increments of ~0.003. Please note that the value of 0.822 (and by extension the increment of ~0.003) was edited following the factual accuracy check stage of this appraisal.
- Utility for 'F-VASI75-89' was varied between 0.908 to 0.935, in increments of ~0.003.
- Utility for 'F-VASI90' was varied between 0.908 to 0.945, in increments of ~0.004.
- The increments between the bounds were determined based on the difference between the bounds divided by nine, meaning that a total of 10 different utility values can be explored.
- The utility values for F-VASI75-89 and F-VASI90+ were varied at the same time, such that a 10x10 table of ICERs using different combinations of utility values could be produced.

The results of this analysis are presented in Table 35 for Base Case 1. In this table, the bottom-left ICER refers to the EAG's preferred utility values (ICER = £303,189 for Base Case 1), whereas the top-right ICER refers to the company's preferred utility values (ICER = £182,586 for Base Case 1). The same table is re-produced for Base Case 2 in Table 36. Ultimately, the multi-way sensitivity analysis demonstrates the extent to which the cost-effectiveness results of the model vary when changing between the company's and EAG's preferred utility values for the different model health states. Given the magnitude of QALYs gain in absolute terms, relatively small changes in utility values can have a profound impact on the ICER.

Table 35: Multi-way sensitivity analysis of utility values (aligned to tentative EAG-preferred base case 1) – ICER for ruxolitinib versus vehicle cream

Utility value for…		Utility value for 'No response'									
…F-VASI75-89	**…F-VASI90**	**0.822**	**0.819**	**0.816**	**0.814**	**0.811**	**0.808**	**0.805**	**0.803**	**0.800**	**0.797**
0.935	**0.945**	£221,221	£216,139	£211,286	£206,646	£202,205	£197,951	£193,873	£189,959	£186,200	**£182,586†**
0.932	**0.941**	£228,072	£222,675	£217,527	£212,612	£207,914	£203,419	£199,114	£194,988	£191,029	£187,228
0.929	**0.937**	£235,361	£229,618	£224,148	£218,932	£213,954	£209,197	£204,647	£200,291	£196,117	£192,112
0.926	**0.933**	£243,132	£237,007	£231,184	£225,640	£220,356	£215,314	£210,497	£205,891	£201,482	£197,258
0.923	**0.928**	£251,432	£244,889	£238,677	£232,772	£227,153	£221,798	£216,691	£211,813	£207,150	£202,687
0.920	**0.924**	£260,320	£253,312	£246,671	£240,370	£234,382	£228,686	£223,260	£218,085	£213,145	£208,424
0.917	**0.920**	£269,859	£262,336	£255,220	£248,480	£242,087	£236,015	£230,240	£224,741	£219,498	£214,494
0.914	**0.916**	£280,124	£272,026	£264,382	£257,157	£250,316	£243,829	£237,670	£231,815	£226,241	£220,929
0.911	**0.912**	£291,201	£282,459	£274,227	£266,461	£259,123	£252,179	£245,596	£239,349	£233,412	£227,762
0.908	**0.908**	**£303,189***	£293,725	£284,834	£276,465	£268,573	£261,120	£254,069	£247,389	£241,052	£235,031

Abbreviations: F-VASI, Facial Vitiligo Area Scoring Index; ICER, incremental cost-effectiveness ratio.

Note: *EAG Base Case 1 – see Table 29; †EAG Base Case 1 with utility values as per Company's preferred assumptions – see Table 32.

Table 36: Multi-way sensitivity analysis of utility values (aligned to tentative EAG-preferred base case 2) – ICER for ruxolitinib versus vehicle cream

Utility value for…		Utility value for 'No response'									
…F-VASI75-89	**…F-VASI90**	**0.822**	**0.819**	**0.816**	**0.814**	**0.811**	**0.808**	**0.805**	**0.803**	**0.800**	**0.797**
0.935	**0.945**	£191,809	£187,403	£183,195	£179,172	£175,321	£171,633	£168,097	£164,703	£161,444	**£158,311†**
0.932	**0.941**	£197,749	£193,070	£188,606	£184,344	£180,271	£176,374	£172,642	£169,064	£165,632	£162,336
0.929	**0.937**	£204,069	£199,089	£194,347	£189,825	£185,508	£181,384	£177,439	£173,662	£170,042	£166,571
0.926	**0.933**	£210,807	£205,497	£200,448	£195,641	£191,059	£186,687	£182,511	£178,517	£174,695	£171,032
0.923	**0.928**	£218,004	£212,330	£206,944	£201,825	£196,952	£192,310	£187,881	£183,652	£179,609	£175,740
0.920	**0.924**	£225,710	£219,634	£213,876	£208,412	£203,221	£198,282	£193,577	£189,090	£184,807	£180,713
0.917	**0.920**	£233,981	£227,457	£221,288	£215,444	£209,901	£204,636	£199,629	£194,861	£190,315	£185,977
0.914	**0.916**	£242,881	£235,859	£229,232	£222,967	£217,036	£211,411	£206,071	£200,995	£196,162	£191,556
0.911	**0.912**	£252,485	£244,906	£237,768	£231,035	£224,672	£218,651	£212,944	£207,527	£202,379	£197,480
0.908	**0.908**	**£262,880***	£254,673	£246,964	£239,708	£232,866	£226,404	£220,290	£214,498	£209,003	£203,783

Abbreviations: F-VASI, Facial Vitiligo Area Scoring Index; ICER, incremental cost-effectiveness ratio.

Note: *EAG Base Case 2 – see Table 31; †EAG Base Case 2 with utility values as per Company's preferred assumptions – see Table 33.