Leniolisib for untreated activated phosphoinositide 3-kinase delta syndrome (APDS) in people 12 years and over [ID6130]

Evidence Assessment Group Report

Produced by Newcastle University

Authors

Nicole O'Connor, Research Assistant, NIHR Innovation Observatory, Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK

Hangjian Wu, Research Associate, Health Economics Group, Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK

Oleta Williams, Research Assistant, NIHR Innovation Observatory, Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK

Cristina Fernandez-Garcia, Health Economics Group, Senior Research Associate, Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK

Sola Akinbolade, Research Associate, NIHR Innovation Observatory, Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK

Ayomikun Ogunyemi, Research Assistant, NIHR Innovation Observatory, Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK

Sonia Garcia Gonzalez-Moral, Research Associate, NIHR Innovation Observatory, Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK

Fiona Beyer, Principal Research Associate, NIHR Innovation Observatory, Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK

Stephen Rice, Senior Lecturer, Health Economics Group, Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK

Nick Meader, Principal Research Associate, NIHR Innovation Observatory, Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK

Correspondence to Nick Meader,
Population Health Sciences Institute,
Newcastle University, Newcastle upon Tyne,
NE2 4BN

Date completed 26 July 2024

Source of funding:	This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR136172.
Declared competing interests of the authors:	None
Contributions of authors:	Nick Meader acted as project lead. Nicole O'Connor and Fiona Beyer acted as lead effectiveness reviewers. Cristina Fernandez-Garcia acted as lead health economist. Sonia Garcia Gonzalez-Moral acted as lead reviewer of the literature search methods. Sola Akinbolade and Ayomikun Ogunyemi acted as effectiveness reviewers. Hangjian Wu acted as a health economics reviewer. Oleta Williams reviewed the literature search methods.
Acknowledgements:	We gratefully acknowledge the expert clinical input from Professor Sophie Hambleton and Professor Andrew Gennery of Newcastle Hospitals NHS Foundation Trust. We also thank Katie Hay, NIHR Innovation Observatory, Newcastle University for assistance with formatting.
Rider on responsibility for the report:	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.
The report should be referenced as:	O'Connor N, Wu H, Williams O, Fernandez-Garcia C, Akinbolade S, Ogunyemi A, Garcia Gonzalez-Moral S, Beyer F, Rice S, Meader N. Leniolisib for untreated activated phosphoinositide 3-kinase delta syndrome (APDS) in people 12 years and over [ID6130]. A Highly Specialised Technology Evaluation. Newcastle upon Tyne: Population Health Sciences Institute, Faculty of Medical Sciences, Newcastle University 2024.
Report key:	[redacted]

Abbreviations

AE	Adverse events
AIC	Akaike's Information Criterion
APDS	Activated PI3K delta syndrome
BI	Budget impact
BIC	Bayesian information criterion
BID	Bis in die (Twice a day)
CE	Cost-effectiveness
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CEM	Cost-effectiveness model
CfB	Change from baseline
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CS	Company's submission
CSR	Clinical study report
DSU	Decision Support Unit
EAG	Evidence Assessment Group
EAP	Early Access Programme
EQ-5D	European Quality of Life-5 Dimensions
FAD	Final appraisal document
FDA	Food and Drug Administration
HR	Hazard ratio
HRQoL	Health-related quality of life
HSCT	Hematopoietic stem cell transplantation
HSUV	Health state utility value
HST	Highly Specialised Technology
HTA	Health technology assessment
IC	Indirect comparison
ICER	Incremental cost-effectiveness ratio
IgM	Immunoglobulin
IPTW	Inverse probability of treatment weighting
IRT	Immunoglobulin replacement therapy
ITC	Indirect treatment comparison
ITT	Intention-to-treat
LY	Life year
MAIC	Matching-adjusted indirect comparison
MCS	Mental component summary score (SF-36)
MeSH	Medical subject headings
MTA	Multiple Technology Appraisal
MTC	Mixed treatment comparison
mTOR	Mammalian target of rapamycin
NA	Not applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NR	Not reported
OS	Overall survival
PAS	Patient access scheme
PFS	Progression-free survival
PfC	Point for Clarification
PCM	Physical component summary (SF-36)
PI3Kδ	Phosphoinositide 3-Kinase delta

PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PtGA	Patient Global Assessment
QALY	Quality adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
RR	Risk ratio
SAE	Serious adverse events
SD	Standard deviation
SE	Standard error
SF-36	Short Form 36
SLR	Systematic literature review
SoC	Standard of care
SPD	Sum of the products of two largest perpendicular diameters of lesions
STA	Single Technology Appraisal
TA	Technology Assessment
TEAE	Treatment emergent adverse events
TTO	Time trade-off
TTOT	Time-to-off treatment
UK	United Kingdom
USA	United States of America
WHO	World Health Organization
WTP	Willingness-to-pay

Table of Contents

Table of Tables

Table of Figures

1 EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence 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 identified by the EAG. Section 1.2 presents the model outcomes. Section 1.3 summarises all key issues identified by the EAG relating to clinical effectiveness and cost-effectiveness. Section 1.4 summarises the EAG's preferred assumptions and ICERs.

Further detail regarding key and non-key issues are described in the main EAG Report (Sections 3 to 7).

All issues identified represent the EAG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the EAG's key issues

Table 1: Summary of EAG's key issues

Issue number	Brief summary of issue	Report section(s)
1	The comparator arm used in the pivotal, phase III, randomised controlled trial (2201 Part II) lacks generalisability to current clinical management in the UK.	Section 4.2.2.1
2	Using a discount rate of 1.5% to the QALY gains.	Section 5.2.2
3	In the probabilistic sensitivity analysis, where uncertainty information was not available, the company assumed standard error to be 10% of its mean for parameters.	Section 5.2.3
4	Assumption that the effect of leniolisib treatment will not wane throughout the lifetime of patients.	Section 5.2.4
5	Additional utility gain from the emotional benefit of leniolisib included in the model.	Section 5.2.6
Abbreviations: EAG = Evidence Assessment Group; QALY = Quality-adjusted life year		

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 the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Reducing the incidence and prevalence of manifestations for APDS, which in turn reduces quality of life decrements experienced by patients receiving leniolisib compared to those under current clinical management.
- Treatment discontinuation leads to increased incidence rates of manifestations associated with APDS, returning to rates experienced under current clinical management. Therefore, treatment discontinuation is an important driver of differences in quality of life.

Overall, the technology is modelled to affect costs by:

- Leniolisib costs: treatment costs for leniolisib with a patient access scheme (PAS) are the biggest contributor to the difference in costs between the leniolisib and current clinical management groups.
- Reduced treatment and monitoring costs of manifestations for APDS: the cost associated with immunoglobulin replacement therapy (IRT), HSCT costs, Immunosuppressant costs and tonsillectomy costs are all higher for patients receiving current clinical management than patients treated with leniolisib (including the cost of leniolisib with a PAS itself)
- Treatment discontinuation: discontinuation from leniolisib treatment reduces costs for leniolisib, increases incidence rates for manifestations (implying increased treatment use associated with various manifestations).

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

- Discount rate: the company used a 1.5% discount rate for health effects and 3.5% discount rate for costs. Using a 3.5% discount rate to both costs and health effects, as per NICE health technology evaluations (HTE) manual,[1] has a large impact on the ICER.
- QALY gains weight: the company applied a decision maker modifier (1.5 QALY) weight to the incremental QALYs accrued by the leniolisib group in the base-case economic analysis. Removing the QALY gains weight has a large impact on the ICER.
- Treatment waning: the company assumed that there is no loss in the efficacy of treatment in the economic model. The impact of treatment waning can be seen as analogous to treatment discontinuation for part of the cohort. The EAG note that the effect of treatment waning on the ICER can be crudely explored by varying the discontinuation rate. The EAG found that incorporating treatment waning through an increase in the discontinuation rate has a large impact on the ICER.
- Additional utility gain from the emotional impact of leniolisib: the company applied an additional 0.1 utility gain to the leniolisib arm to include the benefit from reduced emotional burden due to the lower expected risk of developing various manifestations and mortality. The EAG found that removing this utility gain has a large impact on the ICER.

1.3 Description of the EAG's key clinical and economic issues

Table 2: Uncertainties in clinical evidence relating to Study 2201 Part II (RCT)

Report section	4.2.2.1
Description of issue and why the EAG has identified it as important	The EAG considers the lack of an active comparator arm in 2201 Part II as a key issue in the clinical evidence. The RCT comparator arm received placebo plus selected concomitant treatments. In the UK, immunosuppressive medications, specifically mTOR inhibitors and rituximab, are part of current clinical management for APDS.Patients with prior use of certain immunosuppressive medications were required to complete a protocol-defined washout period to be eligible for enrolment. More importantly, concurrent use of certain immunosuppressive medications was excluded from the clinical trials, due to potential increased risk of infections. This exclusion raises concerns about the generalisability of the comparator. Additionally, baseline imbalances, novelty of the surrogate primary endpoints, and a small sample size used in the analysis introduce uncertainties in the true magnitude of effect, warranting cautious interpretation of the results.
What alternative approach has the EAG suggested?	None
What is the expected effect on the cost effectiveness estimates?	Inputs to the economic model based on the RCT may overestimate the cost-effectiveness of the technology as the comparator group in the study differs from the current clinical management group included in the model.
What additional evidence or analyses might help to resolve this key issue?	The company provided an indirect treatment comparison that used patient data from the ESID registry as a control arm; their care likely better reflects standard care. Therefore concerns about generalisability are partially addressed by this additional evidence.
Abbreviations: EAG = Evidence Assessment Group; APDS = Activated phosphoinositide 3-kinase delta syndrome; mTOR = mammalian target of rapamycin; RCT = randomised controlled trial.	

Table 3: Key issue - 1.5% discount rate applied to future QALYs

Report section	5.2.2.3
Description of issue and why the EAG has identified it as important	The company has applied a discount rate of 1.5% to future QALYs. However, the EAG consider this to be a deviation from the NICE reference case.[2] The EAG believe that this deviation is insufficiently justified and does not meet the NICE reference case criteria: 1) The technology is for people who would otherwise die or have a very severely impaired life; 2) It is likely to restore them to full or near-full health; 3) The benefits are likely to be sustained over a very long period.[1]
What alternative approach has the EAG suggested?	The EAG suggest that the base-case analysis includes a discount rate of 3.5% for both costs and health effects as per the NICE reference case. A sensitivity analysis incorporating a 1.5% discount rate applied to future costs and effects has been included for consideration by the committee.
What is the expected effect on the cost effectiveness estimates?	The use of the reference case discount rate to both costs and effects will increase the ICER reducing the cost-effectiveness of leniolisib compared with current clinical management.
What additional evidence or analyses might help to resolve this key issue?	Longer-term follow up data would provide the necessary evidence needed to assess whether the criteria set by NICE for the application of the 1.5% discount rate are fully met.
Abbreviations: EAG = Evidence Assessment Group; ICER = Incremental Cost-Effectiveness Ratio; QALY = Quality-Adjusted Life Year	

Table 4: Key issue - standard error to be 10% of its mean for parameters where uncertainty information was not available

Report section	5.2.3.2
Description of issue and why the EAG has identified it as important	In the probabilistic sensitivity analysis, the company assumed the standard error (SE) to be 10% of its mean for parameters where uncertainty information was not available, which applies to most of the input parameters in the economic model. A review of previous NICE technology appraisals found that most models typically assumed SEs between 10-30% of the mean (20% was most common) for such parameters.[3] Therefore, in the EAG's view, the rationale for assuming an SE 10% of the mean was insufficiently justified.
What alternative approach has the EAG suggested?	The EAG suggested using a more a conservative value (i.e., an SE that is 20% of the mean) in the probabilistic sensitivity analysis where uncertainty information was unavailable or there was insufficient evidence.
What is the expected effect on the cost effectiveness estimates?	The use of a 20% SE for parameters where uncertainty information was not available implies that a more conservative approach to characterising uncertainty is taken. There is no expectation that the ICER will change in a specific direction.
What additional evidence or analyses might help to resolve this key issue?	In the future, further information about the level of uncertainty around key input parameters (such as treatment effectiveness and utilities) can be obtained through clinical trials, observational data or clinical expert opinion.
Abbreviations: CS = Company Submission; EAG = Evidence Assessment Group; HST = Highly Specialised Technology; ICER = Incremental Cost-Effectiveness Ratio; SE = Standard Error	

Table 5: Key issue - assumption that the effect of leniolisib treatment will not wane throughout the lifetime of patients

Report section	5.2.4.1
Description of issue and why the EAG has identified it as important	The company has stated that the benefits of leniolisib are expected to be sustained during the lifetime of patients. Hence, an assumption of no loss in efficacy has been incorporated in the economic model. This assumption is based on available follow-up data for up to six years. The EAG note that there is no published evidence that the efficacy of leniolisib will continue beyond 6 years.
What alternative approach has the EAG suggested?	The EAG acknowledge there are several different potential approaches, each with their own associated difficulties, that could be taken to incorporate a treatment waning effect into the company's model. The EAG have chosen to adjust the model's discontinuation rate to incorporate the possibility of treatment waning. The EAG propose that the mean discontinuation rate derived from the company's own Expert Consultancy project is applied to the base-case analysis. Further testing of potential efficacy waning was also explored in the EAG sensitivity analyses. The EAG acknowledge that this approach has a significant limitation of excluding the cost of leniolisib treatment.
What is the expected effect on the cost effectiveness estimates?	A reduction in the long-term efficacy of leniolisib has the potential to significantly reduce the QALY difference between groups and decrease treatment costs in the leniolisib group.
What additional evidence or analyses might help to resolve this key issue?	Obtaining longer term data is the best way to establish if the efficacy of leniolisib will be sustained.
Abbreviations: EAG = Evidence Assessment Group; QALY = Quality-Adjusted Life Year	

Table 6: Key issue - additional utility gain from the emotional benefit of leniolisib included in the model

Report section	5.2.6.3
Description of issue and why the EAG has identified it as important	The company assumed that APDS patients receiving leniolisib were expected to benefit from reduced emotional burden due to the lower expected risk of developing various manifestations and reduced mortality, and increased hope due to the availability of a new treatment. Therefore, the company applied a further utility gain, in addition to the assumed utility improvements associated with a reduction in rates of manifestations. The EAG note that this assumption has a large impact on the ICER, however the company has not provided sufficient justification regarding the quantification of this additional impact on utility.
What alternative approach has the EAG suggested?	The EAG suggest removing this assumption given its lack of justification.
What is the expected effect on the cost effectiveness estimates?	Removing the utility gain assumption increases the ICER through a decrease in QALYs for the leniolisib arm.
What additional evidence or analyses might help to resolve this key issue?	Further evidence on the impact of utilities for leniolisib patients, due to a reduction in emotional burden, would help to evaluate the validity of this assumption.
Abbreviations: EAG = Evidence Assessment Group; ICER = Incremental Cost-Effectiveness Ratio; QALY = Quality-Adjusted Life Year	

1.4 Summary of the EAG's preferred assumptions and ICER

Table 7: Summary of EAG's preferred assumptions and ICER

Technologies	**Total costs**	**Total QALYs**	**Incremental costs**	**Incremental QALYs**	**ICER (£/QALY)**
CS base-case – Probabilistic (without QALY gain weight)					
Leniolisib	[redacted]	[redacted]	[redacted]	11.57	[redacted]
SoC	1,613,679	[redacted]			
Fixing errors (1-3) – Probabilistic (without QALY gain weight)					
Leniolisib	[redacted]	[redacted]	[redacted]	11.49	[redacted]
SoC	1,620,167	[redacted]			
EAG base-case – Probabilistic					
Leniolisib	[redacted]	[redacted]	[redacted]	4.51	[redacted]
SoC	1,646,253	[redacted]			
Abbreviations: CS = company submission; EAG = Evidence Assessment Group; ICER = Incremental Cost-Effectiveness Ratio; QALY = Quality-Adjusted Life Year; SoC = Standard of Care					

Table 8: Summary of key EAG scenario analysis results – deterministic analysis: leniolisib versus standard or care

Scenario #	EAG base-case input	Alternative input	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
	EAG base-case	N/A	[redacted]	3.54	[redacted]
1	Treatment discontinuation rate = 14%	Treatment discontinuation rate = 0%	[redacted]	9.49	[redacted]
2		Treatment discontinuation rate = 10%	[redacted]	4.04	[redacted]
3		Treatment discontinuation rate = 30%	[redacted]	2.80	[redacted]
4	Discount rate = 3.5% for both costs and health effects	Discount rate =1.5% for both costs and health effects	[redacted]	4.62	[redacted]
12	No lower limit on utilities	Lower limit on utilities elicited from TTO tasks	[redacted]	3.12	[redacted]

Abbreviations: EAG = Evidence Assessment Group; ICER = Incremental Cost-ffectiveness Ratio; QALY = Quality-Adjusted Life Year; TTO = Time trade-off

2 PLAIN LANGUAGE SUMMARY

What was the question?

Does the drug leniolisib successfully treat APDS in people aged 12 years and over, and does it provide good value for money to the public?

APDS is a rare genetic disorder caused by an overactive enzyme from a genetic mutation. It affects the production of white blood cells. This leads to frequent infections, lung disease (bronchiectasis), inflammatory bowel disease and increased risk of, malignancies such as lymphomas. In the UK, treatments for APDS are decided on a case-by-case basis and treat the symptoms, not the cause. They may include IRT (immunoglobulin replacement therapy), antimicrobial therapies, immunosuppressive treatments, surgical interventions, and, in severe cases, HSCT (stem cell therapy). Currently, no licenced medicines for APDS are approved by marketing regulators in the UK.

Leniolisib is a drug that has been developed to treat the cause of APDS by fixing the overactive enzyme. This allows white blood cells to develop properly and to fight infection more successfully. It is taken as a tablet twice a day.

What did we do?

This project critiques and summarises the manufacturer's evidence on the clinical and cost-effectiveness of leniolisib for treating APDS in people aged 12 years and older. It focuses on safety (adverse events), efficacy, and value for money. Leniolisib has been tested in 38 people with APDS across three clinical trials, including a 12-week randomised controlled trial in 31 people.

What did we find?

The randomised controlled trial reported improvement in levels of white blood cells for fighting infection and reduction in lymphadenopathy (swelling of lymph nodes), but there are some limitations discussed in the report that should be considered when interpreting the results. The manufacturer conducted an extended trial which is ongoing. So far these positive results are maintained and most adverse effects are mild. The rate of infections and levels of antibodies in the blood (which fight infections) were measured to determine how well leniolisib works outside of a clinical trial setting. An 'indirect treatment comparison' compared infection rates and levels of antibodies in the blood of people taking leniolisib from the extended trial compared to people not taking leniolisib who were in a patient registry. There was an improvement in both measures in both the trial and the indirect treatment comparison. The company presented an economic model which suggested that leniolisib could be cost-effective. A lack of available data made it difficult to be confident about cost-effectiveness. The EAG's own analysis suggest that there is uncertainty about the value for money of leniolisib compared to current clinical management.

3 CRITIQUE OF THE COMPANY'S DEFINITION OF DECISION PROBLEM

3.1 Critique of the company definition of the decision problem

A summary of the company's decision problem in relation to the final NICE scope is presented in Table 9 below. A critique of how the company's economic modelling adheres to the NICE reference case is provided in section 5.2.1.

Table 9: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
Population	People with APDS 12 years of age and older	Adults and adolescents with APDS 12 years of age and older	The population is in line with the: participant eligibility criteria for the pivotal leniolisib trials the anticipated licence wording from the Medicines and Healthcare products Regulatory Agency (MHRA), and the population anticipated to receive leniolisib in UK clinical practice	The EAG note a concern with generalisability and a deviation from the decision problem with regards to the starting age in the economic model. Further information is provided in Section 3.1.1.
Intervention	Leniolisib	Leniolisib	N/A – decision problem is aligned with final scope	The intervention is in line with the NICE scope
Comparator(s)	Established clinical management without leniolisib	Established clinical management without leniolisib, specifically covering: antimicrobials, immunoglobulin replacement therapy (IRT), immunosuppressive therapies (including steroids, rituximab and mammalian target of rapamycin [mTOR] inhibitors), haematopoietic stem cell transplantation (HSCT), surgery	N/A – decision problem is aligned with final scope	Overall, the EAG agrees that the company's choice of comparators is appropriate. However, we note an issue about the choice of the comparator used in 2201 Part II and discuss this further

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
		and other procedures, in line with current practice in the UK.		in Section 4.2.2 below.
Outcomes	Infections Lung function Fatigue Mortality Disease severity Immunophenotype measures (lymphocyte counts, immunoglobulin levels, cytokine and chemokine levels) Immune system function (lymph node size, spleen and liver volume size, use of IRT) Adverse and serious effects of treatment Health-related quality of life (HRQoL)	Immunophenotype measures (including lymphocyte counts [such as naïve B cells], serum immunoglobulin levels, and cytokine and chemokine levels) Immune dysregulation measures (including lymphoproliferation, lymphadenopathy [lymph node size], splenomegaly [spleen volume/size], cytopenias and gastrointestinal manifestations) Immune deficiency measures (infections, use of IRT and antibiotics, and hearing loss) Lung disease (bronchiectasis-associated airway disease and advanced lung disease) Fatigue Malignancy and mortality Disease severity and HRQoL (SF-36 and PtGA) Adverse and serious effects of treatment	All outcomes requested by NICE in the final scope are presented in the evidence submission. Neither lung disease nor mortality were investigated as pre-specified efficacy outcomes in the clinical trial programme for leniolisib. However, safety data are available from the clinical trials for both outcomes, including reports of respiratory disorders, infective exacerbations of bronchiectasis, and deaths. [4,5] In addition, real-world evidence is available for the impact of leniolisib on lung disease. [6,7] This evidence submission addresses lung disease and mortality in Section B.2.6.4 and Section B.2.6.6, respectively.	The EAG considers the outcomes described in the CS to broadly match the final scope issued by NICE. Additional data was provided for measures not specified in the final NICE scope and these inform the economic model. See section 3.1.2 of the CS for further information.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of	The company states that cost-utility analysis was conducted for leniolisib versus the relevant comparator, established clinical management in the UK. As per	The EAG note that PSS costs were not included in the economic model	Some concerns: The company has confirmed in their (Point for Clarification)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
	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. The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account.	the NICE reference case, cost-effectiveness was expressed in terms of incremental cost per quality-adjusted life year (QALY), and costs were be considered from the perspective of the NHS and Personal Social Services. A lifetime horizon was used to capture all costs and benefits associated with leniolisib and relevant comparators		response that the cost-effectiveness analysis presented in the CS adopted a NHS perspective (and not a Personal Social Services perspective).
Subgroups to be considered	None	No subgroups have been considered	N/A In line with NICE final scope	Appropriate
Special considerations including issues related to equity or equality	N/A	N/A	There are currently no licensed treatments available for APDS in the UK. This may lead to sub-optimal and inconsistent use of off-label medicines and variable polypharmacy approaches in the management of APDS.[8] , [9-11]	

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
			Additionally, individuals of African descent are often faced with inequalities in access to HSCT, due to having the lowest probability of finding an appropriately matched unrelated donor. [12] Access to HSCT may also be restricted for some young people with APDS due to the lack of parental consent. [13]	

Source: CS Table 1 (Section B.1.1) [14]

Abbreviations: APDS = Activated phosphoinositide 3-kinase delta syndrome; CS = company submission; EAG = Evidence Assessment Group; HSCT = Hematopoietic stem cell transplantation; NICE = National Institute of Health and Care Excellence; PSS = personal social services

3.1.1 Population

The eligible population defined in the CS includes adults and adolescents with APDS who are 12 years and older.[14] The anticipated marketing authorisation for leniolisib by the MHRA is expected to provide a dosing recommendation only for patients weighing 45 kg or more. However, the EAG highlights a potential issue with generalisability; the British National Formulary states the mean value weight for 12-year old adolescents is 39kg.[15] In the leniolisib clinical trials, as part of the inclusion criteria, participants were required to weigh over 45 kg. Therefore, the anticipated dosing recommendation for patients weighing 45 kg or more may exclude otherwise eligible adolescents.The company clarified that two ongoing paediatric clinical trials (NCT05438407 and NCT05693129) are evaluating leniolisib at lower doses (20–70 mg bid for patients aged 6–11 years, and 10–50 mg bid for patients aged 1–6 years, respectively).[16,17] These studies will support future applications to extend marketing authorisation for younger individuals with APDS, including recommendations based on body weight.[14]

The company's economic model, described in Section B.3.2.2.2 of the CS was run for a cohort starting treatment at age 15. The company justified this deviation by stating that 15 is the median age of people with APDS in the ESID registry. However, since the population in the NICE scope is 12 and over, the EAG considers this to be a deviation from the final NICE scope. This is also inconsistent with the starting age in a company submitted document.[35] See sections 5.2.1, 5.2.2, 7.1.2 below for more information by the EAG.

3.1.2 Outcomes

The EAG considers the outcomes described in the CS to broadly match the final scope issued by NICE. Additional data was provided for antibiotic use and hearing loss as part of immune deficiency measures and for cytopenias and gastrointestinal manifestations as part of immune dysregulation measures these measures inform the economic model. Lung function measures were provided alongside measures of lung disease (bronchiectasis-associated airway disease and advanced lung disease) these also inform the economic model.

3.1.3 Economic Analysis

The CS described the QoL impacts of APDS (Section B.1.4.2), highlighting how the condition limits the patients' ability to continue with work, education and daily living activities. The company stated that the analysis adopted a National Health Service (NHS) and Personal Social Services perspective (Section B.3) in alignment with the NICE final scope. However, no social care costs (i.e, home support, community services, health visitors etc) appear to be included in the model. The EAG note that the analysis adopted a partial perspective (NHS only), excluding any PSS related costs, and considers this a deviation from the final NICE scope. The company confirmed in their response to points for clarification that only a NHS perspective had been adopted.

4 CLINICAL EFFECTIVENESS

This chapter presents a summary and critique of the clinical effectiveness evidence contained within the CS for the treatment of APDS with leniolisib. Section 4.1 provides a critique of the company's systematic review. Section 4.2 provides a summary of the clinical effectiveness and safety results together with a critique of the included studies. Section 4.3 provides a summary and critique of the indirect treatment comparison, and section 4.4 provides the conclusions of the clinical effectiveness section.

4.1 Critique of the methods of review(s)

The company undertook a systematic literature review (SLR) in November 2021, with two updates in May 2023 and April 2024, to identify all clinical evidence on efficacy and safety outcomes associated with leniolisib or other PI3K inhibitors, and current clinical management for the treatment of patients with APDS type 1 and 2. The methods for the company's SLR of clinical evidence are detailed in the CS and Appendix D.[14]

A summary of the EAG's critique is presented in Table 10 below. The EAG's assessments (detailed in bold) are on a three-point Likert scale (key issue, some concerns or appropriate).

Table 10: Summary of the EAG's critique of the methods implemented by the company to conduct the systematic literature review

Systematic review stage	Section in CS where methods are reported	EAG's assessment of the robustness of methods
Data sources	Appendix D1.1, p.9	**Appropriate** An appropriate range of bibliographic databases were used. Additionally, grey literature was searched together with a wide range of conference proceedings and clinical trials registries.
Search strategies	Appendix D1.1, p.9-17	**Some concerns** Search strategies were well reported, although previously indexed subject headings were omitted from all versions of the search and a few abbreviated keywords were found to be missing. As such, it cannot be definitively said that all relevant records were retrieved in the search. There are differences between the original 2021 search and the 2023 update and 2024 targeted search. This may be minor; however, it wasn't explained within the report why the changes were made.
Search filters	N/A	**Not applicable** No search filters were used.
Eligibility criteria	Appendix D1.2	**Some concerns** See section 4.1.2 for further information
Screening	Appendix D1.2	**Some concerns** See section 4.1.3 for further information
Data extraction	Appendix D1.2	**Some concerns** See section 4.1.4 for further information.
Quality appraisal	Appendix D1.2	**Some concerns** See section 4.1.5 for further information.
Abbreviations: CS = company submission; EAG = Evidence Assessment Group; SLR = systematic literature review		

4.1.1 Search strategies

Searches were conducted separately for clinical effectiveness (reported in Appendix D.1.1 pg.8), and for economics (cost effectiveness and cost resource use) and health-related quality of life (Appendix G and I respectively). Searches were appraised by the EAG using the Peer Review of Electronic Search Strategies (PRESS) checklist.[18] Critique of the search strategies for cost effectiveness can be found in section 4.1.1. Searches were conducted from the inception date of databases until 11th November 2021, and updated in May 2023 and again in April 2024, so they can be considered up to date.

4.1.1.1 Sources

The EAG reports that a satisfactory search of conference proceedings was performed in American Academy of Allergy, Asthma & Immunology (AAAAI), Clinical Immunology Society (CIS) North American Conference, American Society of Hematology (ASH), International Primary Immunodeficiencies Congress (IPIC), International Congress of Immunology (IUIS) and European Hematology Association (EHA) to supplement the database searches.

4.1.1.2 Subject headings

Appropriate subject headings were used to reflect the current nomenclature, however previously indexed subject headings were omitted from all versions of the search strategy i.e. PHOSPHATIDYLINOSITOL 3-KINASES (2011); Phosphotransferases (1988-1993); Phosphotransferases (Alcohol Group Acceptor) (1994-1997). They were also omitted from the cost-effectiveness and cost resource use searches (critiqued in section 5.1.1).

Without comprehensive testing, it is difficult for the EAG to quantify the effects that all the issues mentioned may have had on search results, but it seems likely the effects would be relatively minor. Overall, the EAG is satisfied that the search for clinical effectiveness studies was conducted appropriately.

4.1.2 Eligibility criteria

The eligibility criteria are said to be pre-defined in a protocol; however, the EAG has not been able to locate a registered protocol for this SLR. The eligibility criteria described in the SLR are generally consistent with the final NICE scope (Table 9) but more broadly defined in terms of population, intervention, outcomes, and study design, aiming for comprehensiveness. The company initially included but later de-prioritised case studies and highlighted seventeen outcome measures included during the SLR updates. Several of these measures are not specifically defined in the final NICE scope but inform the economic model (gastrointestinal manifestations, cytopenias, hearing loss, and bronchiectasis).

The EAG does not consider differences in population and interventions significant because broadening the scope increases the likelihood of identifying all relevant studies. However, the EAG has concerns about the introduction of outcome measures as this increases the likelihood of a type 1 error (i.e. inappropriately concluding that an effect is statistically significant). Additionally, given the small patient population in an orphan condition, data from case studies may have provided additional information about adverse events or other important information relating to pre-specified outcome measures, although we note some case studies inform the findings, as reported in B.2.12 of the CS. Despite this, the EAG have minimal concerns about deprioritising case studies and note that it is common in systematic reviews to exclude study designs that provide less reliable evidence, though this should be specified in advance.

4.1.3 Screening

The screening methods follow recommended practices for the conduct of systematic reviews. For the original SLR, two reviewers independently conducted screening, with a third reviewer arbitrating when disagreements could not be resolved, rapid methods were used for the targeted update conducted on the 9th of April. The PRISMA flow chart for the original SLR and both updates are presented in appendix D of the CS;[19] in summary, 30 unique studies of interventional and observational design were included and underwent synthesis; 88 unique case studies were included, but all did not undergo data extraction or synthesis. Of the 118 included studies (138 records) 10 reported data relating to leniolisib. To avoid erroneously excluding eligible articles during the targeted updatein April 2024, it would have been more reliable for the second reviewer to check the eligibility of all the excluded records at abstract and full-text stages or alternatively to have double-screened a minimum of 20% of the records and ensuring high agreement before continuing to single screening. The EAG considered this issue to be of minimal concern.

4.1.4 Data extraction

Data from 30 interventional or observational studies were extracted by one independent reviewer. A second reviewer assessed missing data and verified the extracted data, and a third reviewer provided conflict resolution when required. The company have not reported if a data extraction form was piloted, and there is no list of pre-specified data items, but it

appears data were extracted into the HST submission tables. The EAG is satisfied with the methods used for data extraction.

The manufacturer reference-linked publications reporting data from the same study to recognise that more than one publication may have contributed to the data entry; it is not clear what steps were taken to avoid double counting. In addition, it is not clear if efforts were made to contact the study authors for missing data. However, this is less of an issue for the trials which form the main basis of the clinical evidence (2201 Part I, II and EI) since the manufacturer, as the trial sponsor, presumably had access to all data.

4.1.5 Quality appraisal

Quality appraisal was conducted by one reviewer with verification by a second. The company used the Downs and Black checklist to appraise the 30 included interventional and observational studies.[19,20] Quality appraisal for the 88 included but deprioritised case studies was not done. No overall risk of bias judgement was provided for each of the individual studies assessed, although as reported in section B.2.12 of the CS, findings from case studies have been used to inform the evidence base and it is unclear how these were selected for inclusion in the narrative synthesis. No attempt has been made to integrate risk of bias findings or to consider the overall impact of study quality on the results.

The company conducted a second assessment of three studies in the company's clinical trial programme (Section B.2.2 CS). A phase 3 randomised-controlled clinical trial (2201 Part II; NCT02435173) was assessed at the study level using the minimum criteria for assessing the risk of bias and generalisability in parallel RCTs, as described in the NICE user guide for company evidence submission template [PMG24].[2] A single-arm, phase 2, open-label, non-randomised clinical trial (2201 Part I; NCT02435173) and an open-label non-randomised extension study (2201 EI; NCT02859727) were assessed using an adapted version of the Critical Appraisal Skills Programme (CASP) Cohort study checklist.[21]

The company considered 2201 Part II to be of high quality and at low risk of bias; judgements to signalling questions and associated rationale are presented alongside the EAG verification in Table 12. The EAG has noted some issues relating to the concealment of allocation and baseline imbalances in prognostic factors.

Non-randomised studies are inherently prone to bias, especially selection bias, due to limitations in the study design. Despite this, single-arm studies are often used for rare conditions with small populations and where there are ethical considerations of withholding potentially effective treatments. The company considered both open-label trials 2201 Part I and 2201 E1 to be of high quality and of low risk of bias, judgements to signalling questions and rationale are presented alongside the EAGs verification in Table 11. The EAG agrees the studies are of high quality but suggests a moderate risk of bias due to inherent limitations of these types of study design and uncertainty around estimates.

The EAG considers the method used to conduct quality assessment reasonable, and both tools used are in line with NICE recommendations. However, assessments undertaken by two reviewers independently are considered the most reliable method to avoid mistakes and the introduction of the reviewer's own biases.

Table 11: Quality assessment verification for Study 2201 Part I and 2201 E1 (reproduced in part from CS table 13)

Study 2201 Part I			Study 2201E1	
CS Critical appraisal		**EAG Critical appraisal**	**CS Critical appraisal**	**EAG Critical appraisal**
Was the cohort recruited in an acceptable way?				
Yes/no/unclear	Yes	Yes	Yes	Yes
Justification	Participant selection was established by checking through all eligibility criteria at screening. A relevant record (e.g. checklist) of the eligibility criteria was stored with the source documentation at the study site. As described in Section 4.2.1, the trial population was representative of the wider APDS population.	The investigator ensured that all patients being considered for the study met the eligibility criteria. Patient selection was established by checking through all eligibility criteria at screening. Deviation from any entry criterion excluded a patient from enrolment into the study. The 6 recruited patients met the specified eligibility criteria and were representative of the wider APDS population, though none of the patients were below 16 years of age or weighed less than 52kg.	Participants were enrolled from Study 2201; additionally, participants who were treated previously with PI3Kδ inhibitors other than leniolisib could be enrolled if they met the eligibility criteria at the screening visit. As described in Section B.2.3.2, the trial population was representative of the wider APDS population.	Study 2201E1 provided continuation of leniolisib therapy for those who directly enrolled from Study 2201 (Part I and Part II), including participants who had received placebo in Part II or access to leniolisib therapy for individuals with APDS who previously received treatment with PI3Kδ inhibitors other than leniolisib, such as nemiralisib and seletalisib, if they met the eligibility criteria at screening. Eligibility criteria were highly consistent across Study 2201 and Study 2201E1.
Was the exposure accurately measured to minimise bias?				
Yes/no/unclear	Yes	Yes	Yes	Yes
Justification	All participants received leniolisib. Exposure to leniolisib was reported	All patients received the same dose of leniolisib for the same duration. The starting dose was 10 mg followed by 30 mg and 70 mg bid for 4 weeks at each dose level respectively.	All participants received leniolisib. Exposure to leniolisib was reported	An open label, single arm extension trial, all participants received leniolisib. All participants received the same dose and dosing regimen.

Study 2201 Part I			**Study 2201E1**	
Was the outcome accurately measured to minimise bias?				
Yes/no/unclear	Yes	Yes	Yes	Yes
Justification	Commonly used outcome measures were included. Outcome assessments were performed according to a pre-specified visit schedule for all participants. Outcome measures were objective and were performed according to standardised procedures to minimise bias and variability in assessments. The trial was not blinded, but outcome measures were objective.	Sequential blood samples were collected in all patients up to 8 hours after the first dose administration and after the first dose following each escalation to the next dose level. The same imaging modality was used throughout the study for the same patient. MRI or CT imaging of neck, chest, abdomen and pelvis were performed at screen and again post-treatment. Similar measurement methods were used for all patients.	Commonly used outcome measures were included. Outcome assessments were performed according to a pre-specified visit schedule for all participants. Outcome measures were objective and were performed according to standardised procedures to minimise bias and variability in assessments. The trial was not blinded, but outcome measures were objective.	Safety was assessed as a primary endpoint, safety assessment, method for assessments and recording were specified and followed according to schedule of assessment. The occurrence of AE was sought by indirect questioning of participants during study visits, physical examination findings, laboratory test findings or other assessments. AEs were monitored until they resolved or judged to be permanent.
Have the authors identified all important confounding factors?				
Yes/no/unclear	Yes	Yes	Yes	Yes

Study 2201 Part I			**Study 2201E1**	
Justification	Comprehensive baseline characteristics were measured, including demographic and clinical characteristics and prior concomitant medication.	The baseline characteristics of the participants were clearly stated – all data for background and demographic variables were listed by age group and patient. Summary statistics were provided for patients overall. Relevant medical history, current medical conditions, results of laboratory screens, drug tests and any other relevant information were listed by age group and patient.	Comprehensive baseline characteristics were measured, including demographic and clinical characteristics and prior concomitant medication.	Baseline demographic characteristics, and prior concomitant medication use were reported for all patients.
Have the authors taken account of the confounding factors in the design and/or analysis?				
Yes/no/unclear	Yes	Yes	Yes	Yes
Justification	Baseline demographic and clinical characteristics are reported in detail, by age group and participant. As confirmed by expert consultants, differences in baseline characteristics are expected to have minimal impact on the results. Due to the small sample size, no subgroup analyses were performed.	The baseline characteristics of the participants were reported. In agreement with the company, the differences in baseline are not expected to have significant impact on the results.	Baseline demographic and clinical characteristics are reported in detail. Subgroup analyses were performed for participants with prior exposure to leniolisib and placebo, as this may have confounded the results.	There were few notable differences in baseline characteristics between extension study patients who had previous exposure to leniolisib and those with previous exposure to placebo. The company performed subgroup analyses for both groups of patients.

Study 2201 Part I			Study 2201E1	
Was the follow-up of participants complete?				
Yes/no/unclear	Yes	Yes	No – study ongoing	No- study ongoing, expected study completion date is January 2027. NCT02859727
Justification	All participants completed the trial	All participants completed the trial. Post-treatment follow-up was completed for all participants. During the four weeks after the last day of dosing, the patients were followed-up for safety. On Day 112 patients underwent the End of Study visit. None of the patients were withdrawn from the study prematurely.	Long-term data from Study 2201E1 (longest expected data collection period of up to six years and three months) is ongoing.	Study is ongoing
How precise (for example, in terms of confidence interval and p values) are the results?				
Yes/no/unclear	Yes, the results are considered precise	Yes	Yes, the results are considered precise.	Yes
Justification	Patient-level results and measures of variability are provided.	Outcome measures and overall safety and efficacy results were provided for all 6 participants. Summary statistics were provided for all parameters of interest. Measures of variability (for example, confidence intervals and p values) were provided.	Measures of variability (e.g. confidence intervals and p values) are provided. Despite the small sample size, Study 2201E1 observed meaningful within-patient results, some of which reached statistical significance.	The study is ongoing, therefore interim analyses were reported. Some results such as for pharmacokinetic parameters were not included in this interim report. Measures of variability such as confidence intervals and p values were only provided for some outcomes.

Abbreviations: APDS: Activated PI3K delta suCRF: case report form; DMC: data monitoring committee; PD: pharmacodynamic; PK: pharmacokinetic.
Source: Rao et al., 2023,[22] Study 2201 Part II CSR, [23]

Table 12: Quality assessment verification for Study 2201 Part II (reproduced in part from CS table 13)

Study 2201 Part II			
CS Critical appraisal		**EAG Critical appraisal**	
Was randomisation carried out appropriately?			
Yes/no/unclear	Yes	Yes/no/unclear	Yes
Justification	Randomisation numbers were assigned in an ascending, sequential order to eligible participants. The investigator entered the randomisation number on the case report form (CRF). A randomisation list was produced using a validated system that automated the randomisation assignment of treatment arms to randomisation numbers in the specified ratio. This procedure ensured that treatment assignment was unbiased.	Justification	A validated automated system generated the random allocation sequence and assigned eligible participants to treatment and control arms in ascending, sequential order. Novartis Drug Supply Management oversaw the process.[24]
Was the concealment of treatment allocation adequate?			
Yes/no/unclear	Yes	Yes/no/unclear	Unclear

Study 2201 Part II			
Justification	Randomisation data were kept strictly confidential until the time of unblinding and were not accessible by anyone involved in the study, with the following exceptions: data monitoring committee (DMC) members, unblinded pharmacist or authorised designee at the site, unblinded monitor (where used) and the PK bioanalyst. This procedure ensured that treatment allocation was concealed.	Justification	The EAG were unable to identify information to confirm the random sequence allocation was adequately concealed before and until participants were assigned. Specifically, we were unable to locate information regarding the mechanism used for concealment. The EAG identified the following information in the protocol: *"randomisation numbers will be assigned in ascending sequential order"* and *"randomization numbers for part II of the study will be assigned in ascending, sequential order to eligible subjects (see Site Operations Manual for details). The investigator will enter the randomization/treatment number on the CRF"* In addition, the randomisation process had oversight from the Novartis Drug Supply Management chain and "*the randomisation scheme for subjects was planned to be reviewed and approved by a member of the Novartis IIS randomisation Group*".[24] This indicates the process was not done by an external organisation, independent of the enrolment personnel.
Were the groups similar at outset of the study in terms of prognostic factors?			
Yes/no/unclear	Yes	Yes/no/unclear	No

Study 2201 Part II			
Justification	Baseline demographic and clinical characteristics were generally well balanced between the leniolisib and placebo groups. As is common in ultra-rare diseases where trials have small sample sizes, some differences in baseline clinical characteristics between the treatment groups were identified (specifically for history of bronchiectasis and gastrointestinal manifestations) and have been discussed further in Section B2.2.2.	Justification	Clinical characteristics show differences in some prognostic factors between comparator and treatment arms. Specifically, there are substantial differences in bronchiectasis and gastrointestinal disease. In addition, there are smaller differences in multiple other factors in the placebo arm, including history of pneumonia, asthma, herpes simplex, and overall neoplasms benign, malignant and unspecified. All these factors are more prevalent in the placebo group.[14,22] It is difficult to say if the imbalances are indicative of systematic error or have occurred by chance, although we agree with the company that imbalance occurring by chance in very small heterogenous populations is very likely. Nonetheless, the imbalance exists, and because there are substantial imbalances in more than one factor we have judged this question to be 'no'. Potential implications of the imbalances are discussed outside of issues related to the risk of bias in section 4.2.2.2.
Were the care providers, participants and outcome assessors blind to treatment allocation?			
Yes/no/unclear	Yes	Yes/no/unclear	Yes
Justification	Study 2201 Part II was a triple-blinded study: the participants, investigator staff, sponsor persons performing the assessments and data analysts remained blinded to the identity of study treatments. Study drugs were identical in packaging, labelling, schedule of administration, appearance and odour.	Justification	Participants, investigator staff, and sponsor personnel performing the assessments and data analysis were blinded to the identity of participants on study treatment. Study drugs were identical in appearance, odour, packaging, labelling, and schedule. Therefore, bias due to deviations from intended interventions and bias in the measurement of the outcome because of non-blinding is unlikely.
Were there any unexpected imbalances in dropouts between groups?			
Yes/no/unclear	No	Yes/no/unclear	No

Study 2201 Part II			
Justification	No participants withdrew or discontinued treatment prematurely in Study 2201 Part II.	Justification	A participant flow diagram is reported by Rao 2023 and describes participants who were randomly assigned, received intended treatment, and analysed.[22] No participants appear to have discontinued or withdrawn from 2201 Part II, indicating good acceptability and tolerability.
Is there any evidence to suggest that the authors measured more outcomes than they reported?			
Yes/no/unclear	No	Yes/no/unclear	No
Justification	There was no evidence to suggest the authors measured more outcomes than they reported. Conclusions from investigator narratives are drawn, and clearly labelled. Post hoc analyses were conducted on data collected as part of the pre-specified outcomes, and were clearly labelled.	Justification	The company have provided data for a wide range of endpoints and outcomes, some endpoints relate to broadly defined outcome domains/manifestations. Many endpoints have been analyzed in different ways to provide supporting information where primary analysis is limited and to aid the interpretation of clinically meaningful differences, but all are transparently reported. Additionally, the company have provided additional data relating to hepatomegaly upon request. The EAG noted the inclusion of a key primary endpoint (change in naïve B cells out of total B cells) in version 7 of the protocol (July 2017) but considered this of limited concern because it was before the study commenced and before any data collection had taken place (December 2017). The company have clearly reported post-hoc analyses, relating mostly to the identification of clinically meaningful differences.
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?			
Yes/no/unclear	No	Yes/no/unclear	No

Justification	All 31 participants who were randomised to treatment were included in the safety analysis set. For the efficacy analyses, an intention-to-treat analysis was not conducted. The PD analysis set consisted of all participants with any available PD data who received any study drug and experienced no protocol deviations with relevant impact on PD data. However, this is unlikely to have introduced bias into the study results: The first principle of the intention-to-treat analysis is to analyse participants in the intervention groups to which they were randomised, regardless of the interventions they actually received. In Study 2201 Part II, all participants received the intervention to which they were randomised, so this principle is fulfilled by the PD analysis set, and there should be no risk of bias with respect to deviations from intended interventions. As described in Section B2.3.1, four participants were excluded from the PD analysis set due to protocol deviations. Three of these were reported as deviations from inclusion criteria, i.e. participants were actually ineligible for inclusion in the trial, rather than representing post-randomisation exclusions of eligible participants. Therefore, these exclusions should not introduce bias into the results. The deviations from inclusion criteria occurred in line with the 2:1 treatment	Justification	The intervention effect of interest was adherence to the intervention. Four people were excluded from the analysis and reasons provided in the published report by Rao 2023.[22]

Study 2201 Part II			
	allocation ratio (leniolisib: n=2; placebo: n=1), so the benefit of randomisation is likely to have been maintained despite these exclusions. Only one participant excluded from the PD analysis set was eligible for the trial. Therefore, overall, the impact of not conducting an intention-to-treat analysis is expected to be insignificant. Supportive analyses including participants with protocol deviations support results of the main analyses.		

Abbreviations: CRF: case report form; DMC: data monitoring committee; PD: pharmacodynamic; PK: pharmacokinetic.

Source: Rao et al., 2023,[22] Study 2201 Part II CSR, [23]

4.2 Critique of trials of the technology of interest, their analysis and interpretation

The CS includes three clinical studies that examine the efficacy and safety of leniolisib for the treatment of APDS: 2201 Part I, a phase 2 open-label dose-finding study; 2201 Part II, a pivotal RCT; and 2201EI, an open-label extension of Study 2201 Part I and Part II. A total of 38 people received leniolisib across the clinical trial programme.

A summary of trial methodology for the leniolisib clinical trial programme, including sample size, study duration and endpoints, is provided in Table 13.

Table 13: Summary of trial methodology of the leniolisb clinical trials (adapted from Table 6 and section B.2.3.1 of the CS)[14]

	Study 2201 Part I (N=6) (NCT02435173)	**Study 2201 Part II (N=31) (NCT02435173)**	**Study 2201E1 (N=37) (NCT02859727)**
Study design	Phase II, international, multicentre, open-label, non-randomised, within-participant, dose-finding, dose escalation study	Phase III, triple-blinded, randomised, international, multicentre, placebo-controlled study	Open-label, non-randomised, international, multicentre extension study
Study locations	Study 2201 Part I and Part II were conducted at ten sites in nine countries: Belarus, the Czech Republic, Germany, Ireland, Italy, the Netherlands, the Russian Federation, the United Kingdom and the USA		Study 2201E1 was conducted at eight sites in seven countries: Belarus, the Czech Republic, Germany, Italy, the Netherlands, the Russian Federation and the USA
Study duration	12 weeks		Six years and three months
Eligibility criteria	See Table 96 in Appendix N.1.3 for a full list of the eligibility criteria.[19] Participants aged 12-75 years with a APDS-associated pathogenic gene variant in *PI3KCD* or *PIK3R1.* They must exhibit nodal and/or extra nodal lymphoproliferation and at least one measurable nodal lesion on CT/MRI and have a clinical history compatible with APDS. Exclusion criteria include surgical or medical conditions, including HSCT that could affect the pharmacokinetics of leniolisib.	See Table 97 in Appendix N.1.3 for a full list of the eligibility criteria.[19] Participants aged 12-75 years with a APDS-associated pathogenic gene variant in *PI3KCD* or *PIK3R1*. They must exhibit nodal and/or extra nodal lymphoproliferation and at least one measurable nodal lesion on CT/MRI and have a clinical history compatible with APDS. Exclusion criteria include surgical or medical conditions, including HSCT, that could affect the pharmacokinetics of leniolisib.	See Table 98 in Appendix N.1.3 for a full list of the eligibility criteria.[19] Participants aged 12-75 years with a APDS-associated pathogenic gene variant in *PI3KCD* or *PIK3R1*. Exclusion criteria include surgical or medical conditions, including HSCT, that may alter the pharmacokinetics of leniolisib.
Interventions	Three increasing doses of leniolisib (10 mg, 30 mg, 70 mg bid)	2:1 ratio of leniolisib 70 mg or placebo bid	Leniolisib 70 mg bid

Co-primary endpoints	Safety parameters including AEs, physical exam, vital signs, ECG, safety laboratory (haematology, blood chemistry, urinalysis) Dose-PD and PK/PD relationship of leniolisib via single and multiple dose concentrations of leniolisib, and pAkt inhibition in unstimulated and stimulated whole blood	Immunophenotype: CfB in % naïve B cells out of total B cells Lymphadenopathy: CfB in the log10 transformed SPD in up to six of the largest lesions from measurable nodal/lymph node index lesions, selected as per the Cheson methodology from MRI or CT imaging	Safety parameters including AEs, physical exam, vital signs, ECG, safety laboratory (haematology, blood chemistry, urinalysis)
Key secondary endpoints	SF-36, PtGA scores and individual participant narratives	3D volume of index and measurable non-index lesions selected as per the Cheson methodology, and 3D volume and bi-dimensional sizes of spleen SF-36, PtGA scores and individual participant narratives Safety parameters including AEs, physical exam, vital signs, ECG, safety laboratory (haematology, blood chemistry, urinalysis)	Frequencies of infections and other disease complications SF-36, PtGA scores and individual participant narratives
Pre-planned subgroups	N/A	Age group: <18 years and ≥18 years Sex: male and female (added to SAP prior to database lock) Genetic diagnosis: APDS1 and APDS2 (added to SAP post database lock)	N/A

Source: CS (section B.2.3.1)[14]

Abbreviations: 3D = three-dimensional; AE = adverse events; APDS = activated PI3K delta syndrome; bid = Bis In Die (twice daily); bpm = beats per minute; CfB = change from baseline; CT = computed tomography; CYP1A2 = cytochrome P450 1A2; CYP3A = cytochrome P4503A; ECG = electrocardiogram; hCG = human chorionic gonadotropin; MRI = magnetic resonance imaging; mTOR = mammalian target of rapamycin; pAkt - phosphorylated protein kinase B; PD = pharmacodynamic; PI3Kδ = phosphoinositide 3-kinase delta; PI3KCD = phosphoinositide 3-kinase catalytic subunit delta; PI3KR1: phosphoinositide 3-kinase regulatory subunit alpha; PK = pharmacokinetic; PtGA = patient global assessment; SAP = statistical analysis plan; SD = standard deviation; SF-36 = 36-item short form survey; SPD = sum of product of diameters; USA = United States of America; WPAI-CIQ = work productivity and activity impairment plus classroom impairment questionnaire

4.2.1 Study 2201 Part I

A summary of the EAG's critique of the design, conduct and analysis of Study 2201 Part I is presented in Table 14.

Table 14: Summary of EAG's critique on the design, conduct and analysis of Study 2201 Part I trial

Trial design or conduct concept	Section in CS where methods are reported	EAG's assessment
Treatment	B.2.2, Table 6	**Appropriate** The trial comprised a dose escalation phase using three increasing doses of leniolisib (10 mg, 30 mg, 70 mg bid) for a study duration of 12 weeks. All patients received the same dose of leniolisib for the same duration.
Randomisation	NA	Not applicable
Allocation concealment	NA	Not applicable
Eligibility criteria	B.2.2, Table 5	**Appropriate** Eligible participants were adolescents and adults aged 12 to 75 years with a minimum weight of 45kg (for age 12-15) or BMI of 18-35 kg/m^2 (for age 16+) with documented APDS/PASLI (n=6). The EAG agrees that this was in line with the NICE decision problem.
Blinding	NA	Not applicable
Baseline characteristics	Interim Clinical Study Report (Part I) p.46 - 47	**Some concerns** The baseline characteristics of the participants were clearly stated – all data for background and demographic variables were listed by age group and patient. Baseline demographic characteristics were representative of the wider APDS population. However, the youngest participant enrolled was 16 years old and the weight range for the 6 participants was 52.9 – 73.2kg. The EAG have concerns that none of the enrolled participants had the minimum weight of 45kg and age 12 years old.
Dropout rate	B.2.5 Table 14	**Appropriate** No participants withdrew or discontinued treatment prematurely. All participants completed the trial.
Statistical analyses	Interim Clinical Study Report (Part I)	**Appropriate** The primary parameter used as PD marker to select the dose for Part II was % pAkt positive B cells (unstimulated and stimulated samples). A concentration-response model was fitted to link systemic drug concentration and pAkt inhibition at each measured time point. The EAG agrees that this is an appropriate approach.

Outcome measures	Interim Clinical Study Report (Part I) p. 48, 74-75	**Appropriate** Safety, pharmacokinetic and pharmacodynamic parameters were determined in all patients treated with leniolisib. Lymph node sizes, transitional and naïve B cell frequencies (as a proportion out of total B cells) were observed and reported. Effect of leniolisib on PI3Kδ pathway, as assessed by a dose and concentration dependent inhibition of pAkt, was reported.
Results: Efficacy outcomes	Interim Clinical Study Report (Part I) p.54-55, 66-67, 64-65	**Some concerns** PI3Kδ pathway, as assessed by phospho-Akt-positive B cells, was suppressed in a dose and concentration dependent manner over the dose range explored. Lymph node sizes (i.e., sum of products of diameters of pre-identified index lymph nodes) were reduced by 40% and spleen volumes were reduced by 39%. As a proportion of total B cells, a reduction in the frequency of elevated transitional B cells (from 38% to 10%) and an increase of naïve B cell frequency (from 32% to 78%) was observed. There was no appreciable change in liver volume. Assessment of the efficacy of leniolisib to modify health-related quality of life in patients with APDS/PASLI through SF-36 (Short Form 36) Survey and WPAI-CIQ were reported as not conclusive. There were elevations in clinical chemistry parameters of systemic inflammation in APDS/PASLI: High-sensitivity C-reactive protein (hsCRP)/Lactate dehydrogenase (LDH). See section 4.2.1.1 for further discussion of efficacy outcomes.
Results: Adverse events	B.2.10 p112-119; Interim Clinical Study Report (Part I) 12.2 p76-87	**Appropriate** The median duration of leniolisib exposure (11.93 weeks) was in line with the 12 week-duration of the study. There were no deaths or discontinuations reported in Part I of the study. There were no serious adverse events (SAEs) reported during Part I of the study. There were no other significant adverse events reported. No study-drug related AEs were reported. The extension trial reported that leniolisib remained well tolerated throughout a median exposure of 154.71 weeks.[25]
Results: Subgroup analyses		**Not applicable** No subgroup analysis was undertaken

Abbreviations: Akt = Protein kinase B; AUC = area under the plasma concentration-time curve; Cmax = observed maximum plasma concentration following drug administration [mass/volume]; CS = company submission; Ctrough = observed plasma concentration at 12 hours post last dose [mass/volume]; EAG = Evidence Assessment Group; APDS/PASLI = Activated PI3K delta syndrome/ p110δ-activating mutation causing senescent T Cells, lymphadenopathy and immunodeficiency; bid = Bis In Die (twice daily); NICE = National Institute for Health and Care Excellence; pAkt = Phosphorylated Akt; SF-36 = Short Form 36; WPAI-CIQ = Work Productivity and Activity Impairment plus Classroom Impairment Questionnaire; hsCRP/LDH = High-sensitivity C-reactive protein Lactate dehydrogenase; SAEs = serious adverse events; AE = adverse events.

4.2.1.1 Results: Efficacy outcomes

Lymphoproliferation can lead to enlargement of the spleen (splenomegaly) and/or liver (hepatomegaly). For the assessment of the impact of leniolisib on lymphoproliferation, liver and spleen 3D volumes were measured. The company reported that treatment with leniolisib led to no appreciable change in liver volume. In response to the clarification of questions, the company provided a justification for this as "in APDS, it is less common for lymphoproliferation to occur in the liver; lymphoproliferative changes are most commonly seen in the lymph nodes and spleen" (Question A5 PfC p.5).[26] Additionally, the company reported that supplementary analyses of Study 2201 Part II demonstrated a statistically significant reduction with leniolisib compared with placebo in liver bi-dimensional size (p=0.0361) (Question A5 PfC p.6).[26]

In the assessment of the efficacy of leniolisib to modify health-related quality of life in patients with APDS/PASLI through SF-36 (Short Form 36) Survey and Work Productivity and Activity Impairment (WPAI) plus Classroom Impairment (CIQ) Questionnaire (WPAI-CIQ), it was reported that both assessments did not provide conclusive outcomes and that these results may be due to the relatively small sample size, the relatively short evaluation period and the heterogeneity of the patient group including adolescent and adult patients (age range: 16 – 31 years). However, both SF-36 and WPAI-CIQ assessments did not show statistically significant results in the Part II of the study which enrolled a larger sample size.

As part of secondary objectives, measurements were conducted to assess the efficacy of leniolisib in reducing clinical chemistry parameters of systemic inflammation in APDS/PASLI: High-sensitivity C-reactive protein (hsCRP)/Lactate dehydrogenase (LDH). However, of the six patients enrolled in Part I:

- three demonstrated 'isolated' elevations of hsCRP values (9.8 mg/L on Day 84; increase from 6.5 mg/L (on day 35) to 19.8 mg/L on Day EoS; 14.2mg/L on screening). The normal hsCRP range is 0-5 mg/L.
- one patient had a single instance of high LDH value (284 U/L) during the study (on Day 71). The normal LDH range is 100-242 U/L.

4.2.2 Study 2201 Part II

Evidence for the effectiveness of leniolisib in patients with APDS is partly informed by the pivotal phase III study; 2201 Part II (NCT02435173).[4] This randomised controlled trial (RCT) design was robust, using an appropriate randomisation method; patients, investigators and the sponsor were appropriately blinded. However, due to challenges inherent to rare disease populations, the sample size is small (n=31), and the benefits of randomisation are more apparent in much larger samples. Given the condition was characterised in 2013 and the number of people with APDS in England is estimated to be [redacted], the EAG agrees that the number of patients enrolled in the trial is appropriate relative to the overall population.[14] The participants are a separate cohort from those enrolled in the earlier dose-finding 2201 Part 1.

Study 2201 Part II compared the effectiveness of leniolisib against placebo plus selected symptomatic treatments over a 12-week period between December 2017 and August 2021.[23] It was a multi-centre, international trial with 31 patients enrolled across nine countries (United States of America, Belarus, Czechia, Germany, Ireland, Italy, the Netherlands, Russia and Belfast in the UK). Whilst there was one trial site in the UK, the EAG note a limited number of participants from the UK and none from England, which could limit the generalisability of findings to the UK setting. The company provided clarification and evidence to support the study's generalisability (Question A11 PfC).[26] Specifically, clinician confirmation to confirm that the baseline characteristics observed in the study do align with those seen in routine clinical practice and further evidence from the European Society for Immunodeficiencies (ESID) registry and the Early Access Programme (EAP). Despite some variability in clinical characteristics observed in the trial compared to the registry, which could be due to the small sample size the EAG is satisfied with the company's clarification and concur the trial results are likely generalisable to APDS patients seen in England (Question A11 PfC).[26]

A summary of the EAG's critique of the design, conduct and analysis of Study 2201 Part II is presented in Table 15.

Table 15: Summary of EAG's critique of the design, conduct and analysis of Study 2201 Part II

Trial design or conduct concept	Section in CS where methods are reported	EAG's assessment
Eligibility criteria	Appendix N.1.3, Table 97	**Key issue** The eligibility criteria are described briefly in Table 13 above and more comprehensively in Appendix N.1.3 of the CS. The eligibility criteria for the treatment group aligns with the population described in the final NICE scope. However, the EAG does not consider the comparator arm to be representative of UK established clinical management as defined in the NICE decision problem. See section 4.2.2.1 for further information.
Treatment	B.2.3.1, Table 6	**Appropriate** The EAG is satisfied that the intervention is appropriate; a fixed dose of 70 mg of leniolisib was delivered twice daily over a 12-week period.
Randomisation	B.2.3.1, Table 6	**Appropriate** The random sequence was generated using a validated system that automated the random numbers for assignment to treatment arms in the specified ratio (leniolisib (n=21) or placebo (n=10)). The 2:1 ratio was specified in v07 of the protocol (July 2017) before the study commenced. The EAG is satisfied that randomisation was appropriately conducted.[24]
Concealment of treatment allocation	B.2.5, table 13	**Some concerns** The exact procedure used to preserve random allocation was insufficiently described. Version 2 of the protocol states, "*randomization numbers for part II of*

		the study will be assigned in ascending, sequential order to eligible subjects (see Site Operations Manual for details). The investigator will enter the randomization/treatment number on the CRF"[24] See the EAG response to the company's critical appraisal in Table 12, section 4.1.5.
Blinding	B.2.3.1	**Appropriate** This was a triple-blinded study. The participant, investigator, and sponsors were masked to treatment assignment throughout the study, minimising the potential for performance and detection bias.
Baseline characteristics	B.2.3.2, Table 9	**Some concerns** There are some imbalances in baseline clinical characteristics (section 4.2.2.2 below).
Dropout rate	B.2.5, Table 13	**Appropriate** No participants withdrew or discontinued treatment prematurely.
Statistical analyses	B.2.4.2	**Some concerns** Sample size calculations were informed by standard deviations observed in 2201 Part 1 (SD=0.14) for lymphadenopathy. Using a two-sided alpha level of 0.05 a sample size of 30 participants (leniolisib n=20; placebo n=10) was estimated to provide 97% power to detect statistically significant differences. A sample size of 30 was estimated to provide sufficient power (at least 78%) to achieve statistically significant p-values in both co-primary endpoints. Three data sets were analysed: Pharmacokinetic (PK), pharmacodynamic (PD) and safety. The PD data set (All participants with any PD data who received any study drug and no protocol deviations) was used for the analysis of covariance for each co-primary endpoint. A subset of the PD data (B-PD) was used to analyse positive change from baseline (CfB) in percentage of naïve B cells out of total B cells and included only patients with <48% naïve B cells at baseline (n=13). The B-PD data analyses a reduced sample and is therefore underpowered, leading to uncertainties in the magnitude of the effect. However, supportive analysis using the full PD data set is provided and demonstrates a similar trend. For lymphadenopathy, patients with zero nodes at baseline were excluded from the primary analysis. To address the possibility of multiplicity both co-primary endpoints needed to be statistically significant to draw inferences.
Outcome measures	B.2.2, Table 5 Clarification questions. 7-9	**Some concerns** The company provided data for 19 endpoints and outcomes related to the following broad categories: immunophenotype, immune dysregulation, immune deficiency, lung disease, fatigue, malignancy and mortality, HRQoL and adverse events.

		The EAG has some concerns regarding the clinical validity of the co-primary outcome measures (section 4.2.2.4.1).
Results: Efficacy outcomes	B.2.6	**Some concerns** Both co-primary endpoints met statistical significance. See section 4.2.2.4.1 for further details
Results: Adverse events	B.2.10; Clinical Study Report	**Appropriate** The incidence of patients reporting study-drug related AEs was comparable between the two treatment groups. Of those receiving leniolisib, 23.8% (5/21) reported study-drug-related AEs compared to 30% (3/10) in placebo. The majority of patients (74.2%) reported Grade 1 AEs. Eight out of 31 patients had Grade 3 or Grade 4 AEs. All 8 patients who reported study-drug related AEs belonged to genetic diagnosis APDS1; this might be explained by the higher known prevalence of APDS1 and subsequent numbers enrolled in the trial compared to APDS2. All SAEs reported throughout Study 2201 Part II were assessed by the investigator as being unrelated to study treatment. There were no discontinuations in Part II. However, [redacted] pulmonary hypertension approximately 3.5 months after the final dose of study medication; [redacted] was reported as unrelated to the study treatment.
Results: Subgroup analyses	B.2.7; Appendix E	**Appropriate** In relation to the co-primary endpoints an *A priori* planned subgroup analysis was undertaken to assess the impact of age (<18 years and ≥18 years). Post-hoc subgroup analysis to assess difference in sex (male vs female) and genetic diagnosis (APDS1 vs APDS2) was also reported. The results are generally consistent among the overall population. However, it is important to note that findings are exploratory; sample sizes are very small with wide confidence intervals, especially for adolescents in the age assessment (see Appendix E.1.2. Tables 26 – 27).

Abbreviations: AE = adverse event; APDS = Activated PI3K Delta Syndrome; bid = Bis In Die (twice daily); CS = company submission; CfB = change from baseline; EAG = Evidence Assessment Group; HRQoL = health related quality of life; IgM = immunoglobulin; mTOR = mammalian target of rapamycin; PD = pharmacodynamics; PGA = Physician's Global Assessment; PK =Pharmacokinetics; PtGA = Patient's Global Assessment; SF-36 = 36-Item Short Form Survey; SAE = serious adverse event; SD = standard deviation; WPAI-CIQ= Work Productivity and Activity Impairment plus Classroom Impairment Questionnaire

4.2.2.1 Eligibility criteria – comparator arm

Participants in the comparator arm received placebo plus a restricted selection of symptomatic treatments.

Sections B.2.3.1 and B.1.4.3 of the CS, along with expert clinical advice to the EAG, confirm that immunosuppressive medication, specifically mTOR inhibitors, steroids and rituximab, typically form established clinical management for APDS in the UK.

Previous use of certain immunosuppressive medications was prohibited in the clinical trial programme if they were administered within a certain timeframe prior to the first dosing of leniolisib or placebo. To be eligible for enrolment participants who had previously used certain immunosuppressive medications were required to complete a protocol defined washout period (see Table 16 for prohibited medications and corresponding washout criteria). More importantly, concurrent use of some classes of immunosuppressive medications, including mTOR inhibitors (sirolimus, everolimus) and rituximab, which form current clinical management, was prohibited in both the treatment and control arms of Study 2201 Part II. Table 16 provides a non-exhaustive list of immunosuppressive treatments that were prohibited throughout Study 2201 Part II, along with corresponding washout criteria.

The company prohibited the use of immunosuppressives due to evidence suggesting they can lead to an increased risk of infections.[24,27] Four out of five clinicians who participated in expert elicitation indicated they would not combine sirolimus, rituximab, mycophenolate mofetil or cyclosporin with a PI3K inhibitor. Therefore there is justification for their exclusion in a treatment arm, but the EAG considered the exclusion of these treatments for the placebo group to be a substantial limitation. Clinical advisors to the EAG pointed out that APDS patients receiving standard care in the UK may receive these medications. Therefore, the placebo group's treatment regimen was considered to be less intensive than that expected in current clinical practice.

In order to overcome this issue, the company carried out an indirect treatment comparison of leniolisib versus an external control group taken from the ESID Registry who were receiving established clinical care, including immunosuppressive therapies and HSCT (see the EAG critique in section 4.3).

Table 16: Prohibited immunosuppressive co-medications across the clinical trial programmes (2201 Part I, II and EI), reproduced from section B.2.3.1 of the CS

Examples of prohibited immunosuppressive co-medications[a]	Time frame within which co-medication was not permitted
Belimumab Cyclophosphamide	Not permitted within six months prior to first dosing of the study medication
B cell depleting medication (e.g., rituximab)	Not permitted within six months prior to first dosing of the study medication If previously received, absolute B lymphocyte counts in the blood must have regained normal values
Cyclosporine A Mycophenolate 6-mercaptopurine Azathioprine Methotrexate	Not permitted within three months prior to first dosing of the study medication
mTOR inhibitors (e.g. sirolimus, everolimus) Non-selective PI3K inhibitors Selective PI3Kδ inhibitors	Not permitted within six weeks prior to first dosing of the study medication Short-term use for up to a total of five days was allowed but only up to one month prior to enrolment in the study
Glucocorticoids above 10 or 25 mg prednisone or equivalent per day (Study 2201 Part I and Study 2201 Part II/Study 2201E1, respectively)	Not permitted within two weeks prior to first dosing of the study medication

Source: CS Section B.2.3.1
Footnotes: [a]Other immunosuppressive medications where the effects were expected to persist at start of dosing of the study medication were also prohibited.
Abbreviations: mTOR: mammalian target of rapamycin; PI3K: phosphoinositide-3 kinase.

4.2.2.2 Baseline characteristics

Baseline demographics and medication use presented in Table 8 and 10 of the CS are generally comparable across the leniolisib and placebo arms, except from previous sirolimus treatment which was more common in the placebo group (30.0%) compared to the leniolisib group (19.0%).[14]

As noted by the company, there are substantial imbalances in some baseline clinical characteristics (Table 9 of the CS), namely, bronchiectasis and gastrointestinal disorder which are more prevalent in patients randomized to the control arm. Additionally, there are smaller differences in multiple other factors in the placebo arm, including history of pneumonia, asthma, herpes simplex, and overall neoplasms benign, malignant and unspecified. The company suggests imbalances may be due to reporting issues across sites and the limited sample size. While the EAG agrees that balancing baseline characteristics in heterogeneous, ultra-rare populations is difficult, the data nonetheless demonstrates that the participants randomised to the control arm were more severely impacted at baseline compared to participants randomised to the treatment arm, potentially bringing uncertainty into the observed treatment effect for leniolisib. The EAG also note a potential cumulative impact of this, and the issues noted above in section 4.2.2.1 regarding the comparator arm which is less intensive than UK standard of care.

4.2.2.3 Outcome measures

The following measures informed the economic model: lymphadenopathy, splenomegaly, cytopenias, bronchiectasis-associated airway disease, advanced lung disease, malignancy and mortality, infection rates, rate of IRT use and rate of antibiotic use. One endpoint used to inform the economic model, antibiotic use, does not appear to be specified explicitly in the final NICE scope.

The co-primary endpoints used to assess the effect of leniolisib did not inform the economic model. Improvement in immunophenotype was measured using B cell normalisation (positive CfB in percentage of naïve B cells out of total B cells using flow cytometry) and improvement in lymphoproliferation was measured by reduction in lymphadenopathy/ reductions in index lesion size (CfB in the log10 transformed SPD in up to six of the largest lesions from measurable nodal/lymph node index lesions, selected as per the Cheson methodology from MRI or CT imaging).

4.2.2.4 Key Results

4.2.2.4.1 Co-primary endpoints

Both co-primary surrogate endpoints used to measure immunophenotype normalisation and reduction in lymphoproliferation were statistically significant.

For improvement in immunophenotype, the adjusted mean change in naïve B cells as a percentage of total B cells from baseline to Day 85 between leniolisib (n=8) and placebo (n=5) was 37.30% (95% CI: 24.06, 50.54; *P*=.0002). The B-PD data set included patients with <48% naïve B cells at baseline (n=13), thereby reducing the sample size. Supportive analyses using the full PD data set (n=21) are presented in Table 17 below. A modified Delphi with treating clinicians (n=24) determined a ≥20% increase in the percentage of naïve B cells of total B cells after 12 weeks of treatment would be clinically meaningful. In a post hoc analysis of the B-PD analysis set, all patients (n=12) in the leniolisib arm achieved a ≥20% increase compared to none (n=5) in placebo, further information regarding responder analysis is presented in Section B.2.6.1, Table 17 of the CS.[14] The company reports observing increases in the percentage of naïve B cells at each time point throughout OLE up to day 252 (~8 months), see section B.2.6.1 of the CS.

For reduction in lymphadenopathy (i.e., reduction in index lesion size), the log10 transformed SPD of index lesions showed a difference in adjusted mean difference of -0.25, (95% CI: -0.38, -0.12; n=26; *P* = .0006) at week 12. Supportive analysis, which includes all patients from the PD analysis set regardless of the number of lesions at baseline, is provided in Table 17 below. A modified Delphi study with treating clinicians (n=24) described a change of ≥20% (adults) or ≥25% (adolescents) from baseline index lesion SPD as a clinically meaningful change following three months of treatment. Post hoc analysis reported a risk difference of 0.64 (95% CI: 0.16, 0.89; n=27) see Section B2.6.2, Table 19 of the CS for further information relating to responder analysis relating to 2201 Part II.[14] The company reports that after 24-36 weeks of treatment with leniolisib (2201 E1) 24 out of 30 participants (80%) achieved the responder threshold. For adults, this is identified as a reduction from baseline in the index SPD by at least 30% for adults and 45% for adolescents at six months, see section B.2.6.2, Table 20 of the CS for further information regarding responder analysis.

The EAG considers there to be uncertainty about the validity of the surrogate endpoints used, particularly the novel measure of naïve B-cells to total B-cells, in reliably predicting long-term clinical benefits. The company's correlation analysis in 2024 examined the link between surrogate biomarkers and patient outcomes, noting a level 2 evidence association between changes in naïve B-cells and long-term infection rates, as defined by EUnetHTA. Additionally, a 2023 modified Delphi survey retrospectively identified relevant variables and clinically meaningful differences for naïve B cells and lymphadenopathy. Despite some biological plausibility and evidence supporting naïve B cells as an endpoint, evidence of a

consistent association remains unclear. For further details, refer to the company's clarification on lymphadenopathy and immunophenotype (B cell normalisation) (Question A7).[26]

Table 17 Clinical effectiveness data for the co-primary endpoints used in 2201 Part II at Day 85 (Reproduced from Table 16 and 18 of the CS) [14]

	Adjusted mean CfB (SE)[c] for naive B cells Mean change (SE) for index lesions	Comparison of adjusted means			
		Difference	SE	95% CI	2-sided p-value
CfB of percentage of naïve B cell to total B-cells					
Primary efficacy analysis (B-PD analysis set)[a]					
Leniolisib 70 mg bid (n=8)	37.39 (5.35)	37.30	5.74	24.06, 50.54	0.0002
Placebo (n=5)	0.09 (6.66)				
Supportive analysis (PD analysis set)[b]					
Leniolisib 70 mg bid (n=13)	34.70 (5.66)	27.94	6.09	15.02, 40.85	0.0003
Placebo (n=8)	6.76 (5.67)				
CfB in log10 transformed SPD of index lesions					
Primary efficacy analysis (PD analysis set: log10 transformed SPD)[d]					
Leniolisib 70 mg bid (n=18)	−0.27 (0.04)	−0.25	0.06	−0.38, −0.12	0.0006
Placebo (n=8)	−0.02 (0.06)				
Supportive analysis (PD analysis set: sum of square root of the product of diameters)[e]					
Leniolisib 70 mg bid (n=19)	−23.68 (4.17)	−21.91	6.86	−36.12, −7.69	0.0042
Placebo (n=8)	−1.78 (6.11)				

Footnotes: [a]Only included participants in the PD analysis set with fewer than 48% naïve B cells out of total B cells at baseline. [b]Included all participants in the PD analysis set apart from six participants, for the following reasons: one participant did not have a baseline measurement of total B cells; one had no naïve B cells at baseline and did not have post-baseline naïve B cell assessments; and four had naïve B cell percentages of less than 48% at baseline but no assessment was performed at Day 85. [c]Data were analysed using an ANCOVA model with treatment as a fixed effect and baseline characteristics as a covariate. The use of glucocorticoids and concomitant immune replacement therapy at baseline were both included as categorical (Yes/No) covariates. Baseline was defined as the arithmetic mean of the baseline and Day 1 values when both were available, and if either baseline or the Day 1 value were missing, the existing value was used.[d]One participant receiving leniolisib was excluded from the PD analysis set because the baseline index node fully resolved by Day 85, and therefore the "log10 transformed SPD of index lesions" could not be derived. [e]Included all participants from the PD analysis set regardless of the number of lesions at baseline.
Abbreviations: bid: Bis In Die (twice daily); CfB: change from baseline; CI: confidence interval; PD: pharmacodynamics; SE: standard error: SPD: sum of product diameters.
Source: Table 16 and 18 of the CS. [14]

4.2.2.4.2 IgM and Infection

Immunoglobulin levels and rate of infections are secondary endpoints in 2201 Part II and are compared in the indirect treatment comparison to support the real-world effectiveness and

long-term use of leniolisib. Evaluation of serum IgM in 2201 Part II showed a mean decrease of 208.26mg/dL in leniolisib compared to 10.00 mg/dL in placebo from baseline to week 12 and improvement was sustained during 2201 E1 to Day 252 (~8 months) with mean levels falling within the normal reference range as defined by Van Gent et al 2009 and Morbach, 2010,[28,29] (see Figure 11 of the CS[14]). Whilst this sustained reduction in IgM supports improvement in immunophenotype, secondary endpoints are suggestive, and any statistically significant results from post-hoc analysis should be interpreted with caution. Annualised infection rates were lower in participants treated with leniolisib compared to the placebo group (2.690 versus 3.476 infections per year). Please refer to section 4.3 below to see the critique of the indirect treatment comparison.

4.2.2.4.3 Patient relevant outcomes

HRQoL

There are no validated HRQoL measures for APDS. Therefore, HRQoL was assessed using both the SF-36 (v2) and Patient Global Assessment (PtGA) as pre-specified secondary endpoints in 2201 Part II and E1.

No meaningful change from baseline (CfB) was observed in the leniolisib group across any of the eight SF-36 scales in 2201 Part II or I. Additionally, there was no statistically significant difference in the physical or mental component summary scores CfB between treatment arms at week 12 (n=27). When considering long-term follow up data from 2201 E1 there was a mean CfB (SD) at the longest reported time period 208 weeks (~48 months) in three out of eight scales, which exceeded the within-participant meaningful change thresholds for SF-36 domain norm-based scores (see Table 29 of the CS):

- Physical function: mean CfB 5.36 (SD=4.95, n=10).
- General health: mean CfB 9.79 (SD=5.46, n=10) and improvement in this scale was consistent across all timepoints (Week 12, 52, 130, 156).
- Vitality: mean CfB of 10.1 (SD=7.16, n=10).
- Data for the physical and mental component summary measures are not reported for week 208 but data for physical component score (PCS) does show improvement at weeks 12 and 52. PCS mean CfB at 52 weeks was 5.49 (SD=7.28, n=28).

In summary, SF-36 data from 2201 Part II is limited and open-label study data show consistent improvement in general health up to 208 weeks and some improvement in physical component summary measures at both 12 and 52 weeks.

Some improvement in PtGA scores in those receiving leniolisib compared to baseline in 2201 Part II were reported as being within the participant meaningful change threshold (10-20mm) (see section B.2.6.7 of the CS for further explanation). A mean CfB at week 12 for leniolisib was reported as 13.05mm (SD=20.71, n=19) compared to placebo -2.25mm (SD=28.95, n=8, there was no statistically significant difference between the treatment groups (p=0.2113). Long-term mean CfB at Week 208 in Study 2201E1 was 25.63mm points of improvement (SD=26.62, n=10); results for all time points, weeks 12, 182 and 208 were described as generally being greater than the meaningful change threshold of >10mm.

Fatigue

Fatigue (or increased energy levels) was noted as an important outcome by all patients (n=6) in Part I, leading to its inclusion as an exploratory outcome in version 7 (July 2017) of the Part II protocol, measured by tri-axial accelerometer. However, the CS indicates that fatigue was not formally measured as an outcome *per se* during the clinical trial programme and was instead documented through investigator-reported narratives collected at the end of 2201 Part II. Investigator narratives describe positive improvements, with 70% of participants receiving leniolisib reporting increased physical activity tolerance and decreased fatigue

compared to 44.4% receiving placebo. Given the importance highlighted by patients, a direct patient-reported measure of fatigue may have provided more informative data rather than the clinician's impression. Further, the sample size is not reported, and the data capture is potentially subject to recall bias. A published case series with six years of follow-up data reported five out of six participants from Study 2201 E1 experienced increased physical capabilities within six months and improved socialisation within one year of leniolisib treatment,[30] this supports qualitative data provided in section B.2.6.5 of the CS. [14] Additionally, data from the EAP reports that 53.0% of affected individuals had clinically meaningful improvements in chronic fatigue, with 27.0% of affected individuals achieving remission.[31]

4.2.2.4.4 Adverse events

Across the clinical trial programme, the adverse events (AE) and treatment emergent adverse event (TEAEs) reported in participants who were administered leniolisib were classified as follows: 82.0% (433/528) were Grade 1-2 and 10.0% (53/528) were Grade 3-4.

SAEs

In 2201 Part II, 20.0% (2/10) of participants on placebo reported six SAEs, compared to 14.3% (3/21) treated with leniolisib who reported five SAEs. In 2201 E1, 21.6% (8/37) of participants experienced 36 treatment-emergent SAEs. The most frequently affected system organ class for SAEs was infections and infestations (13.5% (5/37)) and gastrointestinal disorders (8.1% (3/37)). The investigator considered all the SAEs reported in 2201 Part I, II and 2201 E1 to be unrelated to leniolisib.

Discontinuation

Across the clinical trial programme, six people (16.2%) discontinued treatment. All were enrolled in 2201 E1 (n=37). Reasons include death, adverse event, physician decision, participant/guardian decision (all n=1), and study termination by sponsor (n=2). See CS for more information. The EAG asked the company to clarify the plausibility of classical Hodgkin Lymphoma in a larger sample size. The company responded that this AE was not considered to be related to leniolisib, by the investigator. Including data from the global EAP with over 200 years of exposure to leniolisib seven individuals with APDS have discontinued treatment. See company response to PfCs for more information.

Deaths

[redacted] deaths occurred in the trials; [redacted] completed 2201 Part II [redacted] prior exposure to leniolisib; [redacted], leading to discontinuation and then death. [redacted] were considered unrelated to the study drug.

4.3 Critique of trials and data identified and included in the indirect comparison

The company carried out an indirect treatment comparison to provide evidence about leniolisib compared to standard care (to add to the evidence from the RCT about leniolisib compared to placebo). The treatment arm included patients in the extension study,[25] and the control arm comprised eligible patients from the ESID study registry who were receiving more representative standard care.

A summary of the EAG's critique of the design, conduct and analysis of Study 2201 Part II is presented in Table 18.

Table 18: Summary of the EAG's critique of the ITC methods

Aspect of analysis design or conduct	Section in CS where methods are reported	EAG's assessment
Statistical methods	Appendix N.2, pg. 244;[19] Whalen (submitted 2024)[32]	**Appropriate** The company conducted a comparative analysis using treatment data from patients in the single-arm extension trial and control data from eligible patients in the ESID registry. This is an appropriate way to compare leniolisib to standard care, where the trial control group contained patients receiving placebo and not standard care. The company carried out inverse probability of treatment weighting (IPTW) analyses to adjust for baseline differences between these samples. This is an appropriate method to use in this context, where there is individual patient data (IPD) available for both arms of the study.
Included study characteristics and demographics	Appendix N.1.3 table 98, pg. 228[19] Appendix N.2.2 pg. 244[19]	**Some concerns** The treatment group was taken from the extension trial, and the control group was taken from registry data. Similar eligibility criteria were applied to both groups except that the age and weight criteria were not applied to the control group. See section 4.3.1 for further details.
Covariates included and excluded in the IPTW	Appendix N.2.2 pg. 245;[19] Whalen (submitted 2024) pg. 8[32]	**Appropriate** For the respiratory infections analysis, the company included age, IRT use at baseline, and baseline infection rate. For the serum IgM analysis, the company included age, sex, baseline serum IgM levels, and APDS mutation status. The EAG asked the company to provide their rationale for the selection of covariates in each analysis and agreed with these rationales.
Weighting of covariates	Whalen (submitted 2024)[32]	**Some concerns** IPTW analyses were not always successful in achieving balance between the treatment and control groups for all baseline characteristics. In addition, some covariates were not included in the IPTW analyses, which may have influenced the outcomes; see section 4.3.2.
Outcomes	Appendix N.2.2 pg. 244-5[19]	**Some concerns** The company analysed i) incidence of respiratory infections and ii) change in serum IgM values. These are clinically appropriate outcomes and included in the NICE decision problem.

		The EAG has some concerns about the follow-up time, which are expanded in section 4.3.3.
Results	Appendix N.2.2. pg. 246;[19] Whalen (submitted 2024) Figure 1, pg. 12; Figure 2, pg. 15[32]	**Some concerns** The company reported a statistically significantly lower rate of respiratory infections in the treatment group, with a rate ratio of 0.34 (95% CI 0.19 to 0.59) for leniolisib versus standard care. The company reported that the treatment group experienced a difference in median annualised change in IgM of -1.09 g/L (95% CI: -1.78 to -0.39, p=0.002) compared to the control group (i.e. leniolisib reduced serum IgM more than standard care). The company stated that the indirect comparison provided results consistent with the RCT.
Sensitivity analyses	Whalen (submitted 2024) pg. 12 + 15[32]	**Appropriate** The company reported sensitivity analyses that explored the impact of incorporating different data and covariates, and reported that the results were consistent across the analyses. The EAG's concerns about the sensitivity analyses are reported in section 4.3.4.
Abbreviations: APDS = activated phosphoinositide 3-kinase delta; CS = company submission; EAG = Evidence Assessment Group; ESID = European Society for Immunodeficiences; HSCT = hematopoietic stem cell transplantation; IgM = immunoglobulin M antibody; IPD = individual patient data; IPTW = inverse probability of treatment weighting; IRT = immunoglobulin replacement therapy; PI3KCδ, PIK3R1 = types of phosphoinositide 3-kinase delta		

4.3.1 Included study characteristics and demographics

For the treatment group of the ITC, participants from part I (dose finding study) and part II (RCT) were eligible plus two further patients: aged 12-75 years with documented APDS-associated pathogenic gene variant in PIK3CD (APDS1) or PIK3R1 (APDS2) and lymphoproliferation.

For the control group, the company included eligible participants from the ESID registry which is the largest registry of individuals with primary immunodeficiencies worldwide. Patients were excluded if they had only one registry visit, or if they received an alternative PI3Kδ inhibitor or HSCT prior to or on the second visit. Participants from the ESID registry did not have the same eligibility criteria applied as patients entering the trials: there was no restriction by age or weight and younger patients were included in the control group (median age 12 years [IQR 7 to 21]). Control patients were not required to have lymphoproliferation at baseline and 17% did not.

4.3.2 IPTW Analyses

4.3.2.1 Respiratory infections analysis

Baseline differences remained after the IPTW analyses for receipt of IRT, serum IgM, baseline infection rate and APDS type.[32] (Table 1, pg. 10). Control group patients were more likely to have received IRT at baseline, which may lead to an underestimate of the effectiveness of leniolisib. Baseline infection rate was higher in the control group, which may lead to an over-estimate of the effectiveness of leniolisib, although the absolute number of

infections was low. Clinical advice to the EAG suggested APDS type was unlikely to be a prognostic factor, therefore baseline differences were of less concern.

Use of steroids was not included in the IPTW analyses and was more common in the treatment population. Given that steroids target immune dysregulation by inhibiting leukocyte activity and proliferation, this could lead to an overestimate on the effectiveness of leniolisib.[32] (Table 2, pg. 11)

Use of mTOR was considered in the IPTW analyses. However, there were no patients on mTOR in the treatment population, which consisted of patients from the clinical studies. In contrast, 37% of control participants were receiving mTOR at baseline, and 44% at follow-up. . Given that mTOR has been reported as effective in treating lymphoproliferation, this could make leniolisib appear less effective. However, since leniolisib is considered to be an alternative the use of these medications, higher rates of mTOR use are expected in the control group. The company did not include steroid use or mTOR use in the sensitivity analyses.

4.3.2.2 Serum IgM analysis

Baseline characteristics were similar between groups after the IPTW analyses were conducted.[32] (Table 3, pg. 13) However, infections at baseline were not included in these analyses, therefore baseline differences remained.[32] (Table 4, pg. 14)

4.3.3 Outcomes

4.3.3.1 Respiratory infections analysis

For the treatment group eligible infections included otitis media, sinusitis, bronchitis, infective exacerbation of bronchiectasis, respiratory tract infection and pneumonia; for the control group they included otitis media, sinusitis, chest infection and pneumonia because this was the infection data reported in ESID.

4.3.3.2 Serum IgM analysis

The serum IgM outcomes were annualised to account for the fact that there was a longer interval between measurements in the control group than the treatment group. As the median interval between the first and last IgM tests recorded in the treatment population was 254 days, results were considered to be the rate of change per year, without adjustment.

4.3.4 Sensitivity analyses

4.3.4.1 Respiratory infections analysis

Sensitivity analyses for the respiratory infections analysis explored the impact of: i) imputing missing values rather than using complete cases only, ii) using data from partI/II versus the extension trial, iii) including serum IgM, sex and APDS type as covariates as well as age and baseline IRT and infections, and iv) not censoring at HSCT.

4.3.4.2 Serum IgM analysis

Sensitivity analyses for the serum IgM analysis explored the impact of: i) different time points of measurement (e.g. first to lowest or first to last instead of first to second measurement), and ii) not censoring for HSCT.

For both analyses, the company reported that the results were consistent across the analyses. They did not report a test for differences although visual inspection of the forest plots supports this assertion.

4.4 Summary of all company evidence for leniolisib

4.4.1 Methods

The company evidence to evaluate the safety and efficacy of leniolisib in people with APDS comprises:

- a 12-week, open-label, dose-finding study (2201 Part I)
- a 12-week, randomised, triple-blind, placebo-controlled phase III trial (2201 Part II)
- an ongoing, open-label extension study (2201E1)
- an indirect treatment comparison which uses extension study participants as the treatment arm and eligible ESID registry patients as the control arm.

In total, 38 people received treatment in the clinical trial programme, and an additional 72 people have received leniolisib via the EAP, including six people from three participating centres in the UK.

Study 2201 Part I (n=6) was conducted over 12 weeks establishing an optimal dose of 70mg bid and confirming safety and pharmacokinetic profiles. The EAG has no major concerns about the conduct of Part I.[8]

Study 2201 Part II (n=31) appears to have been methodologically sound although some areas, such as concealment of allocation, were at unclear risk of bias.[22] The key issue is that the comparator group did not receive established clinical management as understood in the UK and defined in the NICE scope. Instead they received a placebo plus restricted symptomatic management but immunosuppressants were prohibited, which may have over-estimated the apparent effectiveness of leniolisib. Also, there were imbalances in baseline clinical characteristics, which indicate patients randomised to placebo were more severely impacted at baseline, potentially overestimating any treatment effects. Post-hoc identification of clinically meaningful thresholds was undertaken to determine the proportion of responders to leniolisib. It is unclear how reliably the co-primary surrogate outcomes (specifically, proportion of naïve to total B cells) predict long-term clinically relevant outcomes that reflect a benefit to patients, but the company reasoned that variability in clinical outcomes is large and the sample size would have had to be unreasonably large to have clinical outcomes as primary outcomes. The small sample size is appropriate relative to the estimated number of people known to have APDS. The trial was conducted across ten sites in nine countries, including Belfast in the UK and the EAG is satisfied the results are generalisable to patients in England and the UK.

Study 2201E1 is an ongoing open-label, multicentre, single arm extension study over six years and three months, for participants who participated in Part I or Part II, or who were treated previously with other PI3Kδ inhibitors and fitted study eligibility criteria.[25] It aimed evaluate longer term safety, tolerability, efficacy and pharmacokinetic data over six years. The EAG has no major concerns about the conduct of 2201E1.

Finally, the company conducted an indirect treatment comparison to assess external validity. This analysis compared participants from the 2201E1 extension study to eligible control participants from the ESID registry, providing a more generalisable standard care group compared to that in 2201 Part II.[32] However the eligibility criteria for the treatment and control group were not matched as there was no age or weight restriction for control participants from the registry. The EAG has some concerns because the weighting across arms was not successful for the respiratory infections analysis.

4.4.2 Results

Because this is an active area of evolving research, many aspects of the study appear exploratory. While treatment with leniolisib appears to demonstrate improvement in many

parameters, there is a risk measuring many outcomes and endpoints and analysing them in different ways can lead to chance findings.

Overall, leniolisib appears to be generally well tolerated across the programme, with a median exposure of three years. Most of the AEs/TEAEs were grade 1-2. None of the TEAEs that were reported during 2201 or E1 lead to discontinuation (n=6; 2201 E1). In addition, though there were [redacted], investigators determined they were unrelated to the study drug. The EAG is satisfied that these events are fully reported in documentation provided by the company.

The company reported a statistically significant effect on the co-primary surrogate endpoints, indicating immunophenotype normalisation and reduction in lymphoproliferation maintained across all trials to the interim analysis cut-off in the extension trial.

Results from the indirect treatment comparison show improvements in the more clinically relevant outcomes of serum IgM levels and respiratory infection rates consistent with findings from 2201 Part II.

Findings on quality of life from the SF-36 data were limited and, in the EAG's view, did not demonstrate long-term improvement in health-related quality of life, with the exception of the general health scale. Findings from PtGA scores were more favourable, as were participant narratives collected during 2201 Part II. All participants from Part I mentioned fatigue as important to them; in the EAG's view, a more robust measure of fatigue would have provided better patient-relevant data.

5 COST EFFECTIVENESS

5.1 EAG comment on company's review of cost effectiveness evidence

This section pertains mainly to the review of cost effectiveness analysis studies. However, the search section also contains summaries and critiques of other searches related to cost effectiveness presented in the company submission. Therefore, the following section includes searches for the cost effectiveness analysis review, measurement and evaluation of health effects as well as for cost and healthcare resource identification, measurement and valuation.

Table 19 presents an overview of the EAG's critique of the methods used to identify studies for the review of cost-effectiveness.

Table 19: Summary of the EAG's critique of the methods for the review of cost-effectiveness

Aspect of cost-effectiveness SLR	Section in CS where methods are reported	EAG's assessment
Data sources for cost-effectiveness analysis review	Appendix G.1, p 163-173 Appendix D.1.1, p 14-18	**Some concerns** An appropriate range of databases and grey literature sources were used in the original 2021 SLR and the 2023 update but some sources were omitted in the targeted 2024 SLR.
Search strategies	Appendix G.1, p 163-173	**Some concerns** Bibliographic database searches failed to include previous MeSH subject headings used for indexing Phosphatidylinositol 3-Kinases such as Phosphotransferases/ and "Phosphotransferases (Alcohol Group Acceptor)"/ which could have led to missing relevant literature. The 2021 searches present some differences to the text search terms and the MeSH subject headings used in the subsequent 2023 and 2024 updates although this is not considered a major concern due to the method for de-duplication used.
Search Filters	Appendix G.1.1, p 164-171	**Some concerns** A search strategy for identification of health economic studies was used appropriately in relevant databases. The company did not provide information of the origin of the cost-effectiveness filter and the EAG queried this in the PfC letter which the company responded to on PfC response (Section B.1, pg. 36). For an EAG critique of the methods see section 5.1.1.2
Data sources for model input	Appendix P, p 273 - 274	**Some concerns** A targeted search for proxy utility values was performed. The CS did not provide information on the sources searched, date searched, or number of records retrieved. The EAG queried with the company in the PfC letter, and the company responded on the Response PfC letter (pg. 51-52). An overview of the EAG critique on the methods is included in section 5.1.1.3
Eligibility criteria for inclusion of economic evaluations		Appropriate
Eligibility criteria for inclusion of health state utility value studies		Appropriate

Eligibility criteria for inclusion of resource use and cost studies		Appropriate
Abbreviations: CS = company submission; EAG = Evidence Assessment Group		

5.1.1 Search strategies for cost-effectiveness SLR

Searches were conducted separately for economic studies (cost-effectiveness and cost resource use) and HRQoL (Appendix H). Searches were appraised by the EAG using the Peer Review of Electronic Search Strategies (PRESS) checklist.[18] Searches were conducted from the inception date of databases until 11th November 2021, and updated in May 2023, and again in April 2024 so they can be considered up to date. The EAG's critique of HRQoL searches are in Section 5.2.6.1 while the EAG's critique of the way the use of previous MeSH headings were omitted in all searches (clinical effectiveness, economic, and HRQoL) can be found in Section 4.1.1.2.

5.1.1.1 Sources

The company searched a reasonable range of databases and grey sources: MEDLINE, Embase, The Cochrane Library databases (CENTRAL and Database of Systematic Reviews), CRD databases (DARE, NHS EED and HTA Database), EconLit and ScHARRHUD (School of Health and Related Research Health Utilities Database). For the EAG's assessment of conference sources used in the CS, see Section 4.1.1.1.

In the CS (Appendix G, Table 45, p 173-174) the company describe searching additional grey sources (CEA Registry and EQ-5D) for the SLR update in 2023 only. The EAG note that these searches were not reported in the PRISMA flowchart for cost-effectiveness (Appendix G, Figure 2, p 178). Furthermore, the company performed a targeted search in 2024 where it did not search comprehensively all sources used in the original SLR 2021 and update SLR 2023.While justification is provided for some of the sources not used in the 2024 targeted search, the company does not provide justification for not searching EconLit in the SLR 2024. The implications of not searching systematically all sources are mainly the introduction of potential publication bias.

5.1.1.2 Search filters

Searches were restricted to economic studies using a set of search strings (referred here as 'filter') in combination with the main clinical-focussed concepts. The EAG has not been able to identify if the company used a validated and published filter. The EAG queried this with the company in the PfC letter (B1. p 7) and the company responded in the PfC response (B1. pg.36).[26] The EAG understands that the company developed its own cost-effectiveness filter from two established and validated filters used by The Scottish Intercollegiate Guidelines Network (SIGN) and the Canadian Agency for Drugs and Technologies in health (CADTH). The company has not provided a justification for the need of developing a new filter and, in comparison with the existing cited filters, the EAG can only identify very little similarities. The company removed many cost-effectiveness related lines from the existing filter and replaced them with new lines of search terms in controlled vocabulary and free text. The company has not provided a rationale for this change and the EAG is not able to test if this approach would have resulted in a more or less sensitive strategy for the retrieval of relevant studies.

Without comprehensive testing, it is difficult for the EAG to quantify the effects that all the issues mentioned may have had on search results, but it seems likely the effects would be

relatively minor. Overall, the EAG is satisfied that the search for economic studies was conducted appropriately.

5.1.1.3 Proxy searches for model input

In the response to PfC letter (B6. pg. 51-53) the company discloses a two-phase approach to identifying studies for model input values.[26] Phase one maps to the reported searches and number of included studies in CS Appendix O – Targeted search for proxy utility values (pgs. 253-254).[19] However, phase two was only disclosed in the response to PfC after the EAG raised a query about the methods for identification and selection of studies to populate the model. In this phase a different method to identify relevant studies which mainly consists of expert consultation for the selection of proxy conditions validated by existing previous NICE technologies appraisals is introduced.

The company cites that "*To ensure consistency with prior technology assessments, studies cited in previous NICE technology appraisal were given priority for selection as inputs, over some of the data identified in the Phase 1 searches*" (Response to PfC, pg. 52). [26] The company accompanies this explanation with Table 18 (Response to PfC, pg. 52) in which lists all included studies from the second phase.[26]

Furthermore, phase one only used PubMed as the main source for searching and only included open access studies which would have introduced publication bias in the selection of studies for inclusion. Phase two depends on studies included in previously published NICE technology appraisals.

Both methods present limitations in their own rights that could bias the selection of studies. A systematic literature search would have been the preferred method for the identification of current proxy values. Without further testing the EAG is not able to ascertain the implication of the methods used for the overall model input and results.

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

5.2.1 NICE reference case checklist

Table 20 summarises the NICE reference case checklist and the EAG's assessment on the company's submission in relation to their base-case analysis. The EAG's assessments (detailed in bold) are on a three-point Likert scale (key issue, some concerns or appropriate).

Table 20: NICE reference case checklist

Element of health technology assessment	Reference case	EAG comment on company's submission
Defining the decision problem	From the scope: People with activated phosphoinositide 3-kinase delta syndrome 12 years and older.	**Some concerns** The model considers adults 15 years or older. The company justifies this decision as this was the average age of people with APDS in the Level 1 (mandatory) dataset of the ESID registry (November 2023 dataset). The summary ESID information submitted by the company states however that the mean age at registry is 17.7 See section 5.2.2.1 for further details.
Comparators	From the scope: Established clinical management without leniolisib	**Appropriate** After consultation with clinical experts, the treatment of APDS varies considerably as each individual has different needs. This makes it difficult to consider what established clinical management looks like. The approach considered by the company was including a combination of antimicrobials, immunoglobulin replacement therapy (IRT), immunosuppressive therapies (including steroids, rituximab and mammalian target of rapamycin [mTOR] inhibitors), haematopoietic stem cell transplantation (HSCT), surgery and other procedures was considered appropriate by the EAG.

Element of health technology assessment	Reference case	EAG comment on company's submission
Perspective on outcomes	Outcomes measured in the final scope included: Infections • Lung function • Fatigue • Mortality • Disease severity • Immunophenotype measures (lymphocyte counts, immunoglobulin levels, cytokine and chemokine levels) • Immune system function (lymph node size, spleen and liver volume size, use of immunoglobulin replacement therapy) • Adverse and serious effects of treatment • Health-related quality of life.	**Some concerns** Outcomes included in the cost-effectiveness model were: • Mortality • Incidence rates for various manifestations of APDS and treatment use under current clinical management • HRQoL for proxy conditions measured in QALYs The model does not include adverse effects. See section 5.2.5 for further details.
Perspective on costs	NHS and Personal Social Services (PPS) perspective.	**Some concerns** The EAG note that the company has taken an NHS perspective, not an NHS and PPS perspective. The company confirmed this in their PfC response. See section 5.2.2.2 for further details.
Type of economic evaluation	Cost-utility analysis with a fully incremental analysis.	**Appropriate** The company presented a full cost-utility analysis comparing leniolisib with current clinical management
Time horizon	Long enough to reflect all important differences in costs and outcomes between the technologies being compared.	**Appropriate** A lifetime horizon was used for the cost-effectiveness analysis.

Element of health technology assessment	Reference case	EAG comment on company's submission
Synthesis of evidence on health effects	Based on systematic review.	**Appropriate** The company conducted a systematic search on HRQoL (health-related quality of life) with insufficient information about the search filter used. The company then did a search on proxy conditions without revealing details of the databases, they searched, total number of records retrieved and date the search was performed; these details were provided at the clarification stage (Question B6). [26] See section 5.2.6.1 for further details.
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.	**Some concerns** EQ-5D based utilities with UK value set were not always used for the proxy conditions. See section 5.2.6.3 for further details. **Some concerns** The company applied a QALY weight of 1.5 to the discounted Incremental QALYs in the base case analysis. the EAG believes that the application of this weight is not in line with NICE guidelines.[1] See section 5.2.2.4 for further details.

Element of health technology assessment	Reference case	EAG comment on company's submission
Source of data for measurement of health-related quality of life	Reported directly by the patients or carers or both.	**Some concerns** Due to lack of evidence on HRQoL measured for APDS, utility for proxy conditions were used. The use of proxy conditions can provide inaccurate estimates of the impact of APDS on HRQoL and may complicate the utility calculation process in CEM (cost effectiveness model). See section 5.2.6.2 for further details.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population.	**Some concerns** The baseline utility was elicited through clinician valuation survey. The utility values derived from the studies of proxy conditions were not always a representative sample of the UK population. See section 5.2.6.3 for further details.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit, except in specific circumstances.	**Appropriate** There was no indication of unequal weighting given to individuals.
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.	**Some concerns** Costs and resource use mostly sourced from NHS reference costs, [33] BNF[34] and eMIT.[35] The EAG were not able to verify some of these costs which leads to uncertainty surrounding the total costs included in the model. Furthermore, the EAG prioritised eMIT as the preferred source of unit costs whenever possible as per section 4.4 of the NICE HTE manual. [1] See section 5.2.7.1 for further details.

Element of health technology assessment	Reference case	EAG comment on company's submission
Discounting	The same annual rate for both costs and health effects (3.5%)	**Key issue** The company applied a discount rate of 3.5% to costs and 1.5% to QALY gains which the EAG considers a deviation from the NICE reference case [1] See section 5.2.2.3 for further details.
Abbreviation: EAG = Evidence Assessment Group; NICE = National Institute for Health and Care Excellence		

5.2.2 Decision problem

Table 21 summarises the EAG's critique on the decision problem of the model adopted by the company.

Table 21: Summary of EAG's critique on the decision problem

Aspect of model	Section in CS where methods are reported	EAG's assessment
Defining the decision problem and population	Document B.3.3.1, p. 132; Pharming_ESID Registry Analysis Summary Document_NICE_22May2024[36]	**Some concerns** The economic model includes adults 15 years or older. The company justifies this decision as age 15 is the average age of people with APDS in the APDS in the Level 1 (mandatory) dataset of the ESID registry (November 2023 dataset). The summary ESID information submitted by the company states however that the mean age at registry is 17.7. [36] See section 5.2.2.1 for further details.
Perspective	Document B.3.5, p. 162	**Some concerns** The EAG note that the company has taken a partial perspective (NHS only). No Personal Social Services costs have been added to the model. See section 5.2.2.2 for further details.
Time horizon	Document B.3.2.2, p. 127	**Appropriate** Lifetime
Discounting	Document B.3.2.2, p.131	**Key Issue [1]** The company applied a 1.5% discounting for health effects and 3.5% discount for costs, which is inconsistent with recent NICE guidelines.[1] See section 5.2.2.3 for further details.
QALY gain weighting	Document B.3.9.1, p.175	**Some concerns** The company applies a 1.5 QALY weighting in the base-case deterministic results and uses the produced cost-effectiveness results in the conclusion of this submission. EAG thinks the presentation of results is unclear. See section 5.2.2.4 for further details.
Abbreviations: CS = company submission; EAG = Evidence Assessment Group		

5.2.2.1 Defining the decision problem and population

The economic model was run for a cohort of individuals starting treatment at age 15. The company justifies using 15 as this is the average age of people with APDS registered in the Level 1 dataset of the ESID registry (November 2023 dataset).[36] However, this is inconsistent with the information in the summary results of ESID submitted by the company, which states that the mean age at registry is 17.7.[36] Therefore, the EAG could not verify the appropriateness of the starting age assumption given the inconsistent statement between submitted documents. In order to test this assumption and its potential effect on cost-effectiveness, the EAG have included a sensitivity analysis which assumes the starting age for the cohort of individuals to be 18 years old.

5.2.2.2 Perspective

With respect to costs the EAG note that the CS adopted an NHS perspective deviating from the NHS and PSS perspective in the NICE reference case. The company has confirmed in their PfC response that the cost-effectiveness analysis presented in the CS adopted an NHS perspective (and not a PSS perspective). Costs were confined to the use of primary, secondary and tertiary care services associated with the monitoring and treatment of the manifestations associated with APDS, even though the costs falling on PSS were stated as having been included in the analysis. Subsequently, the company confirmed in their PfC response that the cost-effectiveness analysis presented in the CS adopted an NHS perspective only (and not a Personal Social Services perspective). Given the burden of this condition on activities of daily living, educational and employment outcomes, the EAG considers the burden on Personal and Social Services should have been considered.

5.2.2.3 Discounting

In the summary of the cost-effectiveness analysis (Section B.3), the company states that a 1.5% discount rate for the future health effects was used in their base-case analysis as "leniolisib is expected to be prescribed from age 12 years and is expected to provide substantial and sustained benefits to the quality and length of life of people with APDS."[14] The recent NICE HTE manual recommended a 3.5% discount rate for both costs and effects.[1] The manual states three exceptions when alternative discount rates are acceptable, with all criteria needing to be met: 1) The technology is for people who would otherwise die or have a very severely impaired life; 2) It is likely to restore them to full or near-full health; 3) The benefits are likely to be sustained over a very long period.[1] It is in the EAG's view that the technologies does not sufficiently meet these three criteria.

The EAG asked the company to clarify how the above criteria are met for this HST and therefore to provide further justification regarding the use of a 1.5% discount rate for future health effects. The EAG also requested the company to conduct the base-case analysis using the 3.5% discount rate if they think their submission is unable to meet the NICE criteria [1] stated above. The company has provided justification on the use of the 1.5% discount rate and this is included in their PfC response.[26]

Overall, the EAG agree that the effectiveness evidence submitted by the company suggests that leniolisib substantially decreases the rate of manifestations associated with APDS, alleviate the symptom burden on patients. The company's own elicitation exercise suggest that it may subsequently lead to significant improvements in QoL and life expectancy. However, the EAG also note that there is uncertainty in the effectiveness and duration of leniolisib. First, the drug appears not to eliminate manifestations in all patients from the current clinical evidence. In addition, due to the lack of long-term data and the mean age of participants starting treatment (15 years old in the economic model) it remains unclear whether participants would regain full health or near full health. Finally, the lack of longer term data means that uncertainty remains on whether the benefits are likely to be sustained over a very long period. The EAG hence note uncertainty remains on whether the criteria set by NICE for the application of alternative discount rates is jointly and fully met. The EAG therefore recommend that a 3.5% discount rate is applied to both costs and effects in the base-case analysis. However, the EAG note that the decision whether to apply an alternative discount rate of 1.5% is the responsibility of the committee. Therefore, the EAG applied a 3.5% discount rate to both costs and effects in EAG's own base-case analysis and conducted a sensitivity analysis using the 1.5% discount rate for the both the costs and health effects.

5.2.2.4 QALY gain weighting

The company applied a 1.5 weight for the QALY gain to the deterministic and probabilistic analyses based on the undiscounted QALY gains derived from their economic analysis comparing leniolisib with current clinical management. The EAG requested the company to

provide further clarification regarding whether the application of the QALY weight to the company's base-case analysis is consistent with the NICE HTE manual.

The NICE HTE manual indicates that "*For highly specialised technologies, the committee will consider the size of the incremental QALY gain in relation to the additional weight that would need to be assigned to the QALY benefits for the cost effectiveness of the technology to fall within the highly specialised technologies £100,000 cost per QALY level"* (section sections 6.2.23, NICE HTE manual[1]. The NICE HTE manual further defines the qualifying criteria for the potential application of a QALY gain weight saying that "*For this weight to be applied, there will need to be compelling evidence that the treatment offers significant QALY gains. Depending on the number of QALYs gained over the lifetime of patients, when comparing the new technology with its relevant comparator(s), the committee will apply a weight between 1 and 3, using equal increments, for a range between 10 and 30 QALYs gained" (section sections 6.2.24, NICE HTE manual[1]).*

The EAG acknowledge that the company's undiscounted base-case QALY gains are sizeable and adhere to the criteria set by NICE for highly specialised technologies. The EAG also agrees with the company that the current NICE manual does not comment on whether undiscounted or discounted QALYs should be used to calculate the QALY weight. However, following recent NICE HST evaluations, the use of undiscounted incremental QALYs to calculate QALY weights seems appropriate.

Nevertheless, the EAG believe that the calculated QALY weight should not have been applied to the company's base-case analyses. Based on the NICE HTE manual, it is up to the committee to discuss the weight it attaches to the results of a non-reference case analysis. The committee should then consider whether the application of the suggested QALY weight is appropriate.[1]. The EAG recommend presenting the unweighted CE results alongside with a recommendation of the QALY gain weight that could be applied if the incremental undiscounted QALY gain is bigger than 10.

The EAG recommendation is supported by recently submitted HSTs (e.g., HST30[37]) where QALY weight results have been submitted as part of the submitted exploratory analysis as well as by past NICE guidance on the evaluation of highly specialised technologies that states that *The Committee will consider the size of the incremental QALY gain in relation to the additional weight that would need to be assigned to the QALY benefits for the cost-effectiveness of the technology to fall within the HST £100,000 QALY limit. Although the NICE website indicates that this past document* only applies to appraisals that started before 1 February 2022, the EAG note that no change in guidance specifically applicable to highly specialised technologies has been subsequently published by NICE.

5.2.3 Model structure

Table 22 summarises the EAG's critique on the model structure and inputs adopted by the company.

Table 22: Summary of EAG's critique on the design of the economic model

Aspect of model	Section in CS where methods are reported	EAG's assessment
Type of model	Document B.3.2.2, p.127	**Appropriate** The EAG notes that a limited number of UK- based clinical experts were involved in the model (pathway) validation process, yet acknowledges the availability of clinical experts in this rare field and the structured expert elicitation process the company has organised. See section 5.2.3.1 for further details.
Health states/events and transitions	Document B.3.2.2, p.126-130	**Appropriate** The model has three states: alive on leniolisib treatment; alive not on leniolisib treatment and death. Within each alive health state, the manifestation incidence and prevalence and treatment utilisation were estimated using a partitioned approach.
Modelling Uncertainty	Document B.3.10.1, p.177	**Key Issue [2]** In the probabilistic sensitivity analysis, the company assumed standard error to be 10% of its mean for parameters where uncertainty information was not available. However, the assumed level of uncertainty sits at the lower bound of many HTA studies[26] without justification. See section 5.2.3.2 for further details.
Survival analysis and extrapolation methods	Document B.3.3.3, p.137; Document B.3.3.4, p.149-150;	**Appropriate** Overall survival in the current clinical management arm was informed by a Weibull curve fitted to KM data from the patient-level data obtained from systematic literature review.[38] Clinician's views were used to validate the resultant survival curve. Impact of treatment on survival (represented by a hazard ratio) in the leniolisib arm was informed by clinicians' views given that the effect of leniolisib on mortality was not assessed in the trial. Please see section 5.2.4.2 for further details of EAG'S comments on the impact of treatment on survival
Abbreviations: CS = company submission; EAG = Evidence Assessment Group		

5.2.3.1 Justifications on the type of model structure

The company mentioned that "the final model structure was chosen to reflect key characteristics of APDS and data availability, and was validated by three HTA experts and one UK clinical expert." [14] In the Expert Consultancy (Exercise 3), another three UK clinicians were presented with a diagram containing all of the treatments used in the company's economic model, and agreed that the diagram represents their region's treatment patterns for patients with APDS. [39] The company clarification response included a summary of the expert elicitation process, with details provided, for example, about the identification of

topics and post interview follow-up. The EAG understands that virtual, semi-structured interviews were conducted to validate a number of assumptions including potential modelling approaches/assumptions. The company confirmed the attendance and the contribution of the UK clinical expert in the interview. Agreement was reached on assumptions related to the model structure, which includes (a) modelling the APDS patient population as a whole (not differentiating between APDS1 and APDS2 patients); (b) use of age-dependent cohort level Markov model so it allows the model to track the average age-dependent onset of multiple, key manifestations in line with the progressive nature of APDS; (c) accounting for combinations of manifestations across multiple organ systems in one patient, by modelling each manifestation separately and using an aggregation approach in the calculation of disutility caused by manifestations.

Overall, the EAG acknowledge that the availability of clinical experts in this field would have been limited and that the chosen model structure was justified using an organised elicitation process. Other efforts have been done to justify the use of the current model structure. For example, a SLR has been conducted to search economic evaluations studies in APDS but no published economic models were found. Another search was conducted to identify relevant evidence in other disease areas, and similar approaches for modelling multiple manifestations. Alternative model structures and their drawbacks were discussed in the CS (Section 3.2.2, p.127-128).[14]The EAG therefore considers the company's choice of model structure appropriate.

5.2.3.2 Modelling uncertainty

In the probabilistic sensitivity analysis section (Doc B B.3.10.1, p177), the company stated that "Where empirical probability distributions were not available, the standard error was assumed to take a value equal to 10% of that of the mean."[14] The EAG note that most of the key model parameters for utility, costs and hazard ratios (HR) for the incidence rates of manifestations used a 10% standard error to represent variation in the precision of model parameters in the model.[19] The EAG requested the company to justify the use of the 10% SE in the PSA. The company provided evidence from a review of all full NICE single TAs published in 2013-2014.[3] The review focused on the assessment of the appropriateness of the PSA conducted in the identified TAs and whether the approaches adopted conformed to the relevant guidelines.[3] The company cited one of the findings of this review that "The variation for the parameters was in most cases assumed and not informed by data, with 68% of TAs including at least one parameter where the standard error was assumed to be 10–30% of the mean, with 20% being the most common assumption."[3]

The EAG acknowledge that the company's 10% SE assumption is within the range of the 10-30% from the review. However, this doesn't justify (a) the company's choice of 10%, which is the lower bound of the range, which implies a high level of precision (and therefore certainty) about the estimate, this level of precision does not appear appropriate given that these estimates were not based on directly relevant empirical evidence (the company used clinicians' judgment and proxy conditions to inform key data inputs for the model); (b) the large proportion of parameters this assumption applied to; (c) for input parameters sourced from the clinical experts, the company did not report whether they had checked that a 10% SE adequately covers the uncertainty in the expert estimates. The EAG acknowledge that the company conducted a scenario analysis in which a 20% SE was applied for parameter inputs without available information on uncertainty in the PfC.[26] The results showed that the probability of cost-effectiveness at a willingness-to-pay threshold of £100,000/QALY dropped from [redacted]% to [redacted]%, implying a moderate impact of change of scenario on CE results. The company also conducted a one-way sensitivity analysis (OWSA) in which the range of input parameters were reconstructed using a 20% SE. The results showed some changes in the Top 10 influencing parameters. The company concluded that "the magnitude of the assumed uncertainty is of less consequence for this cost-effectiveness analysis."[26]

Overall, The EAG consider a 20% SE to be more appropriate than a 10% SE given the evidence to support the parameter estimates This more conservative approach better reflects the high level of uncertainty around the estimates. The EAG used a 20% SE assumption in their probabilistic analysis and OWSA.

5.2.4 Treatment effectiveness

A summary of the EAG's view on treatment effectiveness and extrapolation is summarised in Table 23.

Table 23: Summary of EAG's critique on the treatment effectiveness

Aspect of model	Section in CS where methods are reported	EAG's assessment
Treatment effectiveness and extrapolation	Document B.3.2.2, p.132	**Key Issue [3]** The company assumed no treatment waning effect for leniolisib as there is no evidence of treatment waning across the evidence base for leniolisib (including the leniolisib trials.[22 8 25] However, the EAG's clinical experts suggested treatment waning was possible whilst noting the lack of long term data to support or refute continued efficacy. See section 5.2.4.1 for further details.
Impact of treatment on mortality	Document B.3.3.4, p.149 Document B.3.10.3, p.180-181	**Some concerns** The impact of leniolisib on survival (represented by a hazard ratio) was elicited from 4 clinicians. The EAG is concerned that some clinical experts who participated in Exercise 1 of the Expert Consultancy may have had limited experience treating APDS patients, given the ultra-rarity of the condition worldwide. **Appropriate** In a scenario analysis, the company used manifestation-specific mortality risk instead of overall survival impact of leniolisib. A calibration was conducted to ensure visual fit of the predicted survival curve to the APDS Kaplan-Meier curve in standard of care. The company elaborated the calibration approach and conducted scenario analyses using a possible range of calibration values in the document responding to EAG's Point of Clarification.[26] Overall, the EAG agrees with the company's approach to calculate manifestation-specific mortality rate. See section 5.2.4.2 for further details.
Impact of treatment on manifestations	Document B.3.3.4, p.141-147	**Some concerns** As noted above, the impact of leniolisib on some manifestations were estimated using a small number of clinical experts **Some concerns** The evidence the company used to calculate the hazard ratios for the impact of leniolisib on incidence rates of manifestations is inconsistent with the hierarchical table (Doc B, Table 41).[14] See section 5.2.4.3 for further details.
Abbreviations: CS = company submission; EAG = Evidence Assessment Group		

5.2.4.1 The assumption of no treatment waning

The company stated that "Due to the mechanism of action (MoA) of leniolisib (as described in Section B.1.2 of the CS), a treatment waning effect was not expected. Furthermore, no evidence of treatment waning has been observed in the leniolisib clinical trials, with up to six years of published data from Study 2201E1 available [25]', and continued treatment for up to two years in the EAP [31]." In addition, an advisory board that convened with six UK experts, included focused discussion on the implications of the MoA of leniolisib, agreed that based on leniolisib's MoA, they did not foresee any likelihood of treatment response diminishing over time. One clinician also clarified that treatment effect waning in this context would be restricted to biologics (e.g. monoclonal antibodies). Therefore, there is no treatment waning effect assumed in the model. However, the EAG's clinical experts pointed out that there are no long term data to support or refute continuing efficacy over time. Therefore, it is uncertain whether the effectiveness of leniolisib will remain constant over the patient's lifetime. The EAG believes that a scenario analysis exploring the possibility of treatment waning is needed.

In the response to PfC document, the company reiterated that [redacted][40] [41] and aside from poor compliance or discontinuation (in a minority of patients) there is no clinical rationale to expect loss of effect. The company then acknowledged that "the only remaining means of lost effect are discontinuation or poor adherence." The company also claimed that there was no current evidence of poor adherence, as the compliance rates were high in the trials (99%)[42] and in an observational dataset from the US market where the drug (Joenja®) was commercially available (99.3% in terms of time of a year patients complied).[14] Further to the company's response:

- The EAG note that a 3.54% treatment discontinuation has been incorporated in the company's base-case analysis of the economic model. The company stated that this discontinuation rate was sourced from their trial[25] and the EAP survey[31]).
- Having examined the clinical expert in responses to the company's own Expert Consultancy project, the EAG are concerned that the high adherence assumed by the company may be an overestimation. Responses to the survey indicated that the average proportion (%) of patients receiving a PI3Kδ-specific inhibitor expected to discontinue treatment at any point, and for any reason, is 14%. The values provided by the 5 experts surveyed range from 0% to 30%.
- The EAG acknowledge the difficulties of incorporating a potential treatment waning effect into the model, given the available evidence. The EAG have represented treatment waning in the model as total discontinuation from treatment, this has been applied via the incorporation of a higher treatment discontinuation rate based on the company's expert elicitation exercise. The EAG acknowledge that this approach has a significant limitation of excluding the cost of leniolisib treatment. The EAG have tested potential treatment waning by varying the model's discontinuation rate as follows:
    - The EAG have applied a discontinuation rate of 14% to the base-case scenario in line with the mean value derived from the company's expert elicitation exercise[39] to jointly test the impact of treatment adherence and the possibility of discontinuation due to treatment effectiveness waning
    - The EAG have conducted sensitivity analyses using the treatment discontinuation values covering the range elicited from the company's expert elicitation exercise[39] (0%, 10% and 30%).
- The EAG acknowledge that an alternative viable method to incorporate treatment waning would involve assuming that treatment waning does not necessarily lead to treatment discontinuation with people still on leniolisib despite a decrease in efficacy.

This would involve varying the effectiveness across the model cycles. Due to time constraints, this approach has not been tested by the EAG.

- The EAG note that the company has not explained how treatment discontinuation may be linked to the development of manifestations. The EAG would welcome an exploration of the relationship between treatment discontinuation and development of manifestations.

5.2.4.2 Impact of treatment on mortality

- The impact of leniolisib on survival rates

The company used clinicians' opinions to estimate the effects of leniolisib on survival, given that mortality was not assessed in the corresponding trial study.[22] Four clinical experts were asked to provide upper and lower plausible estimates for mortality at specified ages (ages 20 and 40 years) based on initiation of leniolisib at age 12 years. The experts' cumulative hazard median estimate (annualised) under current clinical management was 0.0118. Expert commentary suggested the survival curve for leniolisib should be closer to that of the general population; therefore, each expert's upper plausible estimate for long-term survival on leniolisib treatment after age 12 was used to calculate a cumulative hazard, which was then annualised, resulting in a value of [redacted] This resulted in a mean HR of [redacted] for long-term survival (with SE assumed to be 20% of the mean, i.e. [redacted]). In results validation, the company acknowledged that the annualised cumulative hazard of mortality under current clinical management using real world data (Pharming case series data) is 0.0241, which is higher than estimates based on experts' opinions. However, for consistency in the calculation and to provide a conservative estimate of survival gains for leniolisib, the company decided to adopt the upper plausible estimates of the clinical experts.

The EAG were initially not able to access the results of each individual clinicians' estimates from the modified SEE exercise and thus requested the company to provide the full detail of the survey results, which the company provided together with the PfC responses.[26] Overall, the EAG is concerned that estimates derived from a small number of experts may be subject to high levels of uncertainty. As noted by the company, the number of experts participated in the evaluation exercise is fewer than the minimum recommended by the York Centre for Health Economics reference protocol for expert elicitation.[43] Following this point, the EAG is also concerned that some clinical experts who participated in the exercise may have had limited experience treating APDS patients given the rareness of the disease worldwide, a limitation noted by the company.[39] Experience of treating patients with leniolisib is expected to be even more limited.

The EAG are concerned that the company used expert opinion in calculating annual hazard under current clinical management when real world data is available, However, the EAG acknowledged that this provides a more conservative HR estimate. In addition, the company decided to use the upper plausible range of the mortality estimates under leniolisib treatment in calculating HR estimate as one expert suggested the survival curve for leniolisib should be closer to that of the general population. The EAG are concerned that this assumption can be biased as this is only based on the opinion of one clinical expert (four experts in total participated in the exercise). The EAG are also concerned about the assumed uncertainty underpinning the HR estimate in the PSA. The standard error is assumed to be 20% of its mean without further justification.

- Modelling the effect of leniolisib on survival using manifestation-specific morality risk

The company conducted a scenario analysis (Doc B, p.185-186) in which mortality under current clinical management was calculated through manifestation-specific mortality risks

rather than an overall mortality risk informed by a case series data of people with APDS. The mortality risks associated with each manifestation were obtained from a retrospective analysis study using CVID patients.[44] Table 24 presents the HRs of mortality for each manifestation.

Table 24: HRs of mortality for each manifestation

Manifestations	HR[a]	Source
Lymphoproliferation	1.67	Odnoletkova et al. (2018)[44]
Gastrointestinal manifestations	0.97	
Lymphoma	5.48	
Cytopenia	1.08	
Bronchiectasis	0.83	
Advanced lung disease[b]	4.85	
Infections	1	Assumption
Hearing loss	1	Assumption

Footnotes: [a]HRs represent a comparison of mortality risks with and without the manifestation within the CVID cohort e.g., mortality risk for CVID patients with lymphoma compared to CVID patients without lymphoma.
[b]Reported for granulomatous lymphocytic interstitial lung disease (GLILD).
Abbreviations: HR: hazard ratio.

A calibration factor was applied to ensure the predicted outcomes were in line with observed overall mortality in APDS patients, yet little information was provided on how this calibration value (i.e., [redacted]) was determined.[14] The EAG requested the company to provide full details on the calculation process of this calibration value. In the PfC, the company mentioned that "the calibration factor was a simple static multiplier of the hazard of mortality".[26] They stressed that the calibration was supported by HTA experts.[26] The company also included visualised results of the survival curve when calibration was not implemented, and results of sensitivity analysis in which the calibration value varies from [redacted]. Based on these results, the company justified that the survival curve with a calibration value of [redacted] provides the best visual fit to the observed overall mortality in APDS patients. In general, the EAG agrees with the company's alternative approach of calculating mortality rate employed in the scenario analysis.

5.2.4.3 Impact of leniolisib treatment on manifestations

- Small sample size in the expert elicitation exercise

In the CS's base case analysis, HRs for the impact of leniolisib on the incidence, and proportions of severity reduction and resolution of manifestations were based on a variety of evidence sources, including: the leniolisib clinical trial programmes (Study 2201 Part II, 2201E1 and EAP survey[22 25 31]) and the modified SEE (Exercise 1 of the Expert Consultancy[39]) as outlined in Table 41, Section B.3.3.4 of the company submission.[14] Evidence from the leniolisib clinical trials was given the highest priority, followed by the EAP data, with the modified SEE data used to address any subsequent gaps.[26]

Five clinicians participated in the modified SEE exercise which was used to generate HR estimates, reduction in severity and resolution of manifestations. They were asked to provide their estimates for the upper and lower plausible limits of manifestation occurrence under leniolisib treatment, and the midpoints of their estimates were used in the primary analysis. The EAG is concerned that the estimates elicited from the survey can be subject to a high level of uncertainty given only 5 clinicians participated in the survey.

- The inconsistent use of evidence across groups with different data quality

The EAG also noticed the inconsistency of using evidence informing the HRs for the impact of leniolisib. Although the criteria for the hierarchical use of clinical evidence were set in Table 41 of the CS Section B.3.3.4, the company did not strictly follow these criteria. The company sometimes used clinical opinions when higher quality of evidence was available (e.g., the HR for Cytopenia and for hearing loss). However, the case of HR for hearing loss seems to be special, as the EAG agrees that in the absence of sufficiently reported incidence in the trials and EAP survey, the use of expert opinion may be more appropriate. In addition, the HR value elicited from clinical experts leads to a more conservative estimation of the benefits associated with leniolisib.

5.2.5 Adverse events

Table 25: Summary of EAG's critique on the adverse events within the economic model

Aspect of model	Section in CS where methods are reported	EAG's assessment
Adverse events used within the model		**Some concerns** No comparisons on AE/TEAEs between leniolisib treatment and standard care. See section 5.2.5.1 for further details.
Abbreviations: CS = company submission; EAG = Evidence Assessment Group		

5.2.5.1 Adverse events used within the model

The company did not incorporate adverse effects into the CEM, as most of the AE/TEAEs reported in the trials were Grade 1 or Grade 2 and therefore assumed to have limited impact on HRQoL. The EAG acknowledge that the incidence of reported AE/TEAEs in the studies conducted by the company (see Table 36, CS Doc B) is similar in the leniolisib and placebo groups. However, the EAG notes that participants in the control (placebo) group were required to refrain from using medications (such as mTOR inhibitors, rituximab and cyclophosphamide therapies) commonly used to manage immune dysregulation in this population. Therefore, treatment outcomes in the placebo group may differ from those patients under current clinical management.

The EAG note that there is substantial uncertainty regarding differences in the incidence of AE/TEAEs between the leniolisib and current clinical management arm. The company did not explore alternative assumptions regarding AEs in the model.

5.2.6 Health-related quality of life

5.2.6.1 Searches for health-related quality of life SLR

The company conducted separate searches for the HRQoL SLR. A reasonable range of databases were searched: Embase, MEDLINE, The Cochrane library databases (CENTRAL and Database of Systematic Reviews), CRD Database (DARE, NHS EED and HTA Database), EconLit and ScHARRHUD database. The searches were run from database inception to date of search. The company run the first systematic search on 11th of November 2021, a subsequent update on 18th May 2023 and a targeted update on 23rd April 2024. Some databases were not included in the last targeted update. The company provides justification for not searching some databases which are reasonable, however a rationale for not searching Econlit has not been provided. For the EAG's evaluation of grey literature

sources and conference sources see Section 4.1.1.1. The EAG's critique of the way past subject headings were not included in all searches (clinical effectiveness, economic, and health-related quality of life) can be found in Section 4.1.1.2. The company restricted searches to HRQoL studies using a filter for health utilities/quality of life. The EAG requested further clarification on the providence of the filter in the PfC letter (B.6 pg. 9). The company response provided in-depth clarification on how the filter was developed and the reasons for broadening existing validated and published filters.[26]

Overall, the EAG are satisfied that the search for HRQoL studies was conducted appropriately.

Table 26: Summary of EAG's critique on HRQoL

Aspect of model	Section in CS where methods are reported	EAG's assessment
Identification of HRQoL data within the SLR	Appendix H1.1, p.182-192	**Some concerns** An appropriate range of databases were used, though one relevant database was omitted in the 2024 targeted update, which may have reduced the amount of eligible HRQoL studies identified by the searches. See Section 5.2.6.1 for further details.
Source of preference data for valuation of changes in health-related quality of life	Document B.3.4.1, p.153-154; Document B. 3.4.3, p.155 & Pharming_Exercise 2_EQ-5D-5L HCP Valuation_20Mar2024 [45]	**Some concerns** Based on NICE HTA guidelines, EQ-5D data directly elicited from the patients is recommended to be used in the model.[1] However, SF-36 instead of EQ-5D values were elicited in the trial studies supporting the CS.[22] [8] [25] **Some concerns** Proxy respondents were used to elicit baseline utility value used in the economic model, which can potentially cause bias. See section 5.2.6.2 for further details.
HRQoL evidence used for the cost effectiveness model	Document B.3.4.5, p.162; p.157-161; Document B.3.4.5, p.162 Document B.3.4.5, p.157-161	**Some concerns** EQ-5D measures were not always used in the elicitation of disutility values for the proxy conditions. **Key issue [4]** Insufficient justification for the additional utility gain due to the emotional benefit of leniolisib. **Some concerns** Inconsistent use of source of age-adjusted utility for the general population (Ara&Brazier 2011[46] or Kind et al. 1999[47] for the denominator of manifestation disutility). See section 5.2.6.3 for further details.
The approach of utility calculation		**Some concerns** The additive approach the company used to calculate the overall disutility from the manifestations can be biased.

Abbreviations: CS = company submission; EAG = Evidence Assessment Group; HRQoL = health-related quality of life; SLR = systematic literature review

5.2.6.2 Source of preference data for valuation of changes in health-related quality of life

- No available EQ-5D data elicited directly from the trials

It is expected that EQ-5D data elicited directly from the patients representing the UK APDS population was used in the model based on the NICE reference guidelines.[1] However, SF-36 rather than EQ-5D was measured in the leniolisib clinical trials to evaluate participant HRQoL.[22 8 25] A brief explanation is given as to why SF-36 was initially chosen as the preference measure in the design stage of the trials. However, the company claimed that SF-36 data from the clinical trials could not be used to inform HRQoL in the base case model because (a) SF-36 could not capture the specific HRQoL benefits important for people living with APDS and lacked sensitivity in detecting meaningful changes in certain domains; (b) baseline SF-36 data from the trials have already included the impact of several manifestations of APDS, and therefore would overestimate the impact of APDS. The company conducted a scenario analysis (scenario 7) in which baseline utility was informed by EQ-5D-3L mapped from the SF-36 data.

The company instead used utility values for proxy conditions obtained from various sources (see Table 27). The EAG believes that SF-36 utility estimates derived from the trial may be most applicable to the model population with external validity issues associated with the alternative estimates from other sources. However, the EAG acknowledge that using mapped values from the SF-36 data may not be ideal for this patient groups and no other studies measuring HRQoL directly from APDS patients have been identified.

The EAG have explored the alternative source of utility for each manifestation using values elicited from the clinicians' EQ-5D exercise in the scenario analysis. In a related assumption in the CS' base case analysis, it was assumed that for patients experiencing improvement in severity of manifestations due to leniolisib treatment, the utility decrement due to those manifestations would be reduced by 50%, based on expert opinion. The EAG have conducted a scenario analysis assuming a 25% reduction.

- Using proxy respondents to elicit baseline utility value used in the economic model

The company stated that baseline utility (i.e., [redacted]) was informed by the clinician EQ-5D vignette study in which clinicians were asked to rate a number of health states according to the dimensions of the EQ-5D-5L by assigning the level they perceive is most accurate to represent the patient's HRQoL in their opinion. The EAG is concerned about the utility estimates elicited from proxy respondents, as there is evidence suggesting that proxy participants tend to overestimate impairment and underestimate HRQoL caused by diseases.[48] For this reason, the EAG conducted scenario analyses using either the baseline utility value informed by data from Study 2201 Part II, or the general population utility value calculated by Ara & Brazier (2010).[46]

Summary results of the clinician EQ-5D vignette study was provided, but the EAG initially identified a value different to that used by the company when using the results to produce an EQ-5D estimate a (i.e., [redacted]; sourced from Vignette G, table 5 of the vignette survey report [45]) in the verification process. The company clarified that this utility of [redacted] used in the model was calculated based on the mean APDS general utility values for males and females after mapping to EQ-5D-3L. The EAG encourages the company to improve their presentation and transparency of the results reported in the EQ-5D vignette study.

5.2.6.3 HRQoL evidence used for cost effectiveness model

- EQ-5D based utilities with UK value set were not always used for the proxy conditions

Given the lack of utility data identified from the clinical trials and HRQoL/utility SLR, the company conducted a targeted search to identify utility values associated with APDS manifestations and treatments from proxy conditions. The company also used EQ-5D

surveys completed by the clinicians to derive utility values for the CEM. The utility values, HRQoL methods and proxy conditions for each manifestation were presented in Table 27.

The company aims to include studies using EQ-5D to estimate utility values as much as possible aligning to the NICE guidelines,[1] but acknowledged that "this was not always possible, forming a limitation of the approach." (Section 3.4.5, Doc B, p168[14]) The EAG notices that EQ-5D methods were not used in the calculation of utility multipliers for some manifestations (i.e., Gastrointestinal disorder infection and hearing loss), and non-UK value set was used in the calculation of Cytopenia utility multiplier. In general, the EAG acknowledge the company's attempt in the identification of HRQoL evidence for APDS and accepts the lack of evidence for this rare disease, yet at the same time concerned that these limitations pose challenges to the validity and relevance of the utility values used in the model.

- The assumption of the utility gain from emotional benefit of leniolisib

The EAG understand leniolisib may provide positive impacts on patients' emotional state, therefore affecting HRQoL in addition to the effect captured by the conventional EQ-5D measures. However, no evidence is provided to justify the quantification of this impact in the CS.

The EAG requested the company to provide justification of this utility gain and asked them to conduct a scenario analysis with varying levels of this additional utility gain. In the responses to EAG's request, the company listed three studies in which positive psychological impact, such as positive view of life, optimism and absence of anxiety on HRQoL (measured by EQ-5D), are quantified with results suggesting a utility gain of 0.11-0.17. However, the EAG is still concerned about the validity of this assumption because: (a) the EAG are unsure whether the three observational studies used to justify the assumption of utility gain were identified using a systematic search.[49 50 51 52] It is unclear whether there is evidence of other relevant studies to inform this assumption; (b) the studies identified by the company are based on different cohorts of patients (patients with different conditions in different countries) and therefore the generalisability of the results to APDS patients is uncertain; (c) one of the studies[49] is based on unvalidated study-specific questionnaires, which can lead to biased estimates of the impact of utility gain; (d) perhaps most importantly, as the EQ-5D questionnaire contains a dimension measuring anxiety and depression, there may a double-counting issue if the psychological impact of leniolisib is included in addition to the utility captured by conventional EQ-5D measures.

Regarding the uncertainty of the utility gain in the economic model, the company initially assumed that there was no uncertainty for this utility gain in the PSA. In the PfC, the company conducted a OWSA using a range of 0.08 to 0.12, and a scenario analysis of PSA assuming a 10% standard error around the mean estimate of the utility gain.[26] The company suggested that the results of these additional analyses remain relatively unchanged compared to the initial results in the CS. However, the EAG are concerned that the assumption of 10% standard error is not justified (see Key Issue 2, modelling uncertainty).

Overall, the EAG believe that the evidence used to justify the utility gain due to the psychological impact of leniolisib is highly uncertain and likely to bias CE results. Therefore, the EAG has removed this assumption from the EAG base-case analysis.

- Inconsistent use of source of age-adjusted utility

The company calculated the utility multiplier as the utility for each proxy condition divided by the utility for the UK general population with the same age as the cohort used for the proxy condition. This utility multiplier was then used to calculate the disutility for each manifestation. The EAG note that the age-specific utility for the UK general public used by

the company for the calculation of the utility multiplier for the manifestations was based on the method developed by Kind et al. (1999). [47] In contrast, the company used Ara and Brazier's (2011) method[46] to generate the age-dependent utility decrements. These utility decrements were then used to generate the age-dependent baseline utility applied to the model . The company did not provide justification on the inconsistent use of the methods applied, although the EAG note that this does not have a considerable impact on the CE results.

Table 27: Utility data for proxy conditions used in the economic model

Input description		Utility, disutility or multiplier	Source condition	HRQoL methods
APDS baseline utility (no modelled manifestations or treatments)		[redacted]	APDS	EQ-5D-5L completed by clinicians; Mapped to EQ-5D-3L index scores [5]
Splenomegaly utility multiplier		0.91	Myelofibrosis	EQ-5D-3L (UK value set: Dolan et al., 1997)[53]
Gastrointestinal disorder utility multiplier		0.46	Inflammatory bowel disease (IBD)	Methods unclear (likely a disease-specific measure completed by patients, mapped to EQ-5D, and valued using UK value set by Dolan et al, 1997)[53]
Cytopenia utility multiplier		0.88	Immune thrombocytopenic purpura (ITP)	EQ-5D completed by patients; US value set: Shaw et al. (2005)[54]
Malignancy disutility (first year only)		-0.48	APDS	EQ-5D-5L completed by clinicians; Mapped to EQ-5D-3L index scores (Hernandez Alava et al. (2020)[5]
Malignancy utility multiplier (first year and beyond)		0.86	Diffuse large B-cell lymphoma	EQ-5D-3L completed by patients; UK value set: Dolan et al. (1997)[53]
Infections	**Moderate lower respiratory infection**	-0.003	Not disease-specific	Disability weights based on a global survey (including European countries) that used pairwise comparison methods

	s disutility			in which respondents were asked to indicate which of two health states briefly described to them they considered to be "healthier"
	Severe lower respiratory infections disutility	-0.009		
	Moderate upper respiratory infections disutility	-0.003		
	Herpes zoster disutility	-0.004		
	Infections: weighted average disutility	-0.004		
Bronchiectasis utility multiplier		0.91	Bronchiectasis	EQ-5D-3L completed by patients; UK value set: Dolan et al. (1997)[53]
Advanced lung disease utility multiplier		0.65	Cystic fibrosis	EQ-5D completed by patients; UK value set: MVH group
Hearing loss	**Mild hearing loss disutility**	-0.01	Not disease-specific	Disability weights
	Moderate hearing loss disutility	-0.027		
	Weighted average	-0.02		

Source: CS Doc B, Section 3.4.5 [14]

Abbreviations: CS = Company Submission; EAG = Evidence Assessment Group

5.2.6.4 The approach of utility calculation

The company stated that "starting from the baseline utility, an additive approach was assumed in order to combine the utility impacts of manifestations and treatments when more than one manifestation/treatment is experienced".[14] The EAG thinks that this approach can overestimate the combined effect of disutility when people with APDS experience multiple manifestations. The disutility for people experiencing multiple manifestations can be lower than the aggregated disutilities for each individual manifestations if several similar manifestations affect the same dimension of QoL elicited in EQ-5D questionnaire. The company assumed no lower limit on utility value per cycle. The EAG has explored an alternative assumption in their sensitivity analysis by using a lower limit elicited from the Pharming TTO study. These lower limit value was provided by the company.[6]

5.2.7 Resources and costs

Table 28 summarises the EAG's critique on resources and costs within the economic model.

Table 28: Summary of EAG's critique on resources and costs

Aspect of model	Section in CS where methods are reported	EAG's assessment
Resource use and cost data identified in the SLR	Document B3.1 and Appendix 1	**Appropriate** The EAG agrees that the 3 studies identified by the company on costs and HCRU in APDS did not provide useful evidence relevant to the decision problem for this evaluation.
Intervention costs (Leniolisib acquisition costs)	Document B.3.5.1 and Appendix 1	**Appropriate** The company included information about the full and patient access scheme (PAS) discounted cost associated with the cost per bottle of 60 tablets of leniolisib 70 mg.
Administration costs	Document B.3.15	**Appropriate** The company stated that no additional costs are associated with the administration of leniolisib, beyond acquisition costs.
Adverse event costs	Document B. 3.3.5	**Appropriate**
Health state costs (Manifestations for both Leniolisib and clinical management)	Document B.3.5.2 and appendix K	**Some concerns.** In the absence of published sources of evidence the resource inputs included in the model by the company were based on results from the quantitative survey of the Expert Consultancy project (Exercise 4). The EAG note that there is uncertainty in the ranges of resource use elicited from the experts and this may have an impact on the true level of healthcare resource use applied to both the leniolisib and the current clinical management group. The EAG were not able to verify a substantial number of the unit costs of leniolisib manifestation-specific treatment and current clinical management applied by the company. See section 5.2.7.1 for further comment.
Health state costs (Monitoring for both Leniolisib and clinical management)	Document B.3.5.2 and appendix K	**Some concerns** In the absence of published sources of evidence the resource inputs included in the model by the company were based on results from the quantitative survey of the Expert Consultancy project (Exercise 4). The EAG note that there is uncertainty in the ranges of resource use associated with the monitoring of APDS elicited from the experts and this may have an impact on the true level of healthcare resource use applied to both the leniolisib and the current clinical management group. See section 5.2.7.1 for further comment

Abbreviations: CS = company submission; EAG = Evidence Assessment Group; SLR = systematic literature review

5.2.7.1 Health state costs (Manifestations and monitoring for both leniolisib and clinical management groups)

The company has provided a detailed list of all the manifestation-specific treatment costs for both leniolisib and current clinical management. The EAG notes the following:

- The EAG were not able to verify a substantial number of the unit costs of leniolisib manifestation-specific treatments and current clinical management applied by the company. The EAG suspect that this could be because the company has used outdated eMIT unit prices. The EAG prioritised eMIT as the preferred source of unit costs whenever possible as per section 4.4 of the NICE health technology evaluations manual.[1] The revised list of unit costs included in the EAG base-case analysis has been included in Appendix 1.
- The EAG believe that where the company has included treatment costs for both patients under 18 years of age and over 19 years of age the correct formula for allocating these age-dependent costs has not been applied to the model. The EAG has corrected this issue (as described in section 7.1.1) to ensure that these costs are correctly applied in the excel economic model.
- The company conducted a thorough elicitation process to estimates impact of leniolisib and current clinical management on resource use associated with the manifestations experienced by APDS patients. The EAG acknowledge the difficulties of these process in the field of rare diseases and note that there is uncertainty and large variation in the values elicited from the experts. The EAG note that this may have an impact on the true level of healthcare resource use applied to both the leniolisib and the current clinical management groups.

6 COST EFFECTIVENESS RESULTS

6.1 Company's cost effectiveness results

The company's base-case cost-effectiveness results using the PAS discount are shown in Table 29. The analysis compares the cost effectiveness of patients treated with leniolisib with patients treated with current clinical management for the APDS population. Unweighted (weighted) results suggest that leniolisib increases the health outcomes by 10.46 (15.46) QALYs and increases costs by [redacted] per patient; and being more costly and more effective than the current clinical management pathway (ICER = [redacted] for the unweighted and [redacted] for the weighted).

Table 29: Company base-case cost-effectiveness results (under the PAS discount)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Unweighted results		Weighted results	
						Incremental QALYs	ICER (£/QALY)	Incremental QALYs	ICER (£/QALY)
leniolisib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	10.46	[redacted]	15.46	[redacted]
Current clinical management	1,587,334	34.81	[redacted]						

Sources: CS Doc B, Section 3.9 [14]

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

6.2 Company's sensitivity analyses

6.2.1 Probabilistic sensitivity analysis

To explore uncertainty within their cost-effectiveness analysis, the company conducted a PSA over 1,000 iterations using the PAS price for leniolisib. The company reported the following weighted PSA results showing leniolisib is more effective with, incremental QALYs increasing to 11.57, and more costly ([redacted]) compared with current clinical management.

The company also reported the simulated PSA results for the QALY weighted results showing that leniolisib has a [redacted] probability of being cost-effective compared with current clinical management at a £100,000/QALY WTP threshold. The unweighted results reduced the probability of leniolisib being cost-effective compared to current clinical management to [redacted] at a £100,000/QALY WTP threshold. Table 30 and Figure 1 show the probabilistic results reported by the company.

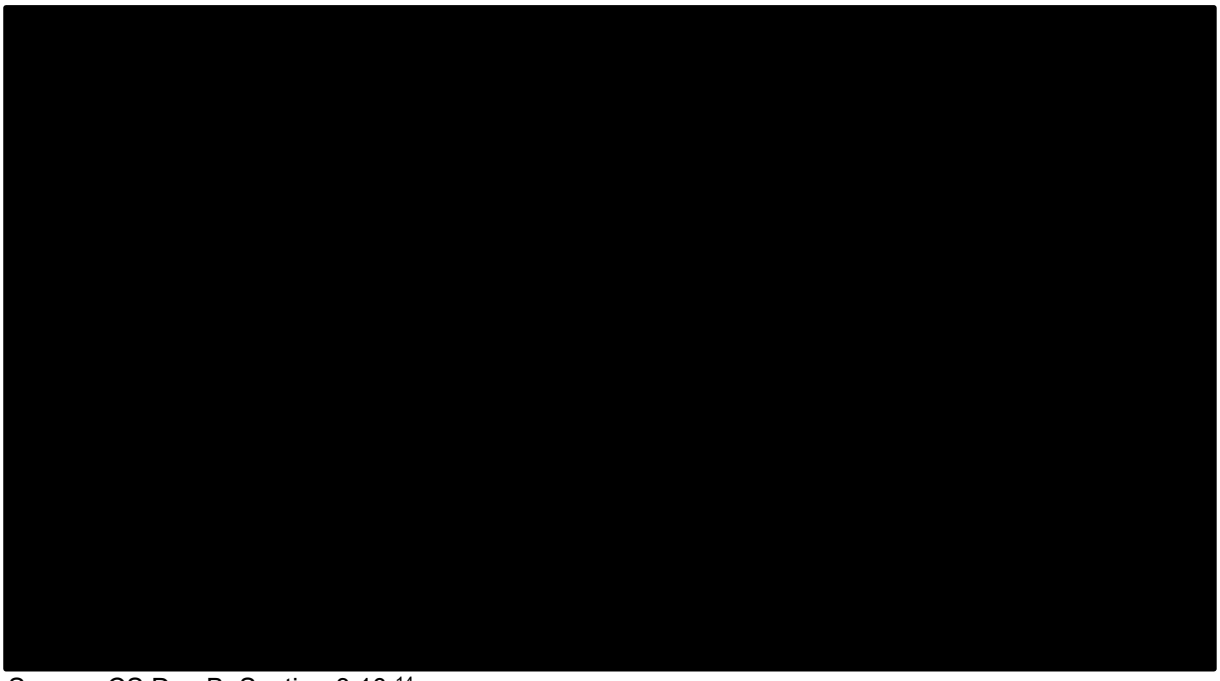
Table 30: Probabilistic base-case results, with QALY weighting (with proposed PAS)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Weighted incremental QALYs	ICER (£/QALY)	Weighted ICER (£/QALY)
leniolisib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	11.57	17.10	[redacted]	[redacted]
Current clinical management	1,613,679	34.77	[redacted]						

Source: CS Doc B, Section 3.9 [14]

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

Figure 1: Scatterplot of probabilistic results



Source: CS Doc B, Section 3.10 [14]

6.2.2 One-way sensitivity analysis

The company also conducted one-way sensitivity analyses (OWSA) by varying a selection of model parameters individually. For parameters where empirically-derived 95% CIs were not available, a SE of 10% of the mean was assumed by the company. Parameters were varied within lower and upper bounds set to 2.5% and 97.5% of their 95% CIs.

The results of the OWSA are presented in Table 31 and Figure 2. The parameters with the greatest influence on the ICER were the rate of gastrointestinal manifestations, the rate of advanced lung disease, and the long-term utility impact of lymphoproliferation and splenomegaly for standard care.

Table 31: OWSA results for leniolisib versus current clinical management (top 10 most sensitive parameters only)

Parameter name	Lower bound ICER (£)	Upper bound ICER (£)	Difference (£)
Age specific manifestation rate of Gastrointestinal manifestations	[redacted]	[redacted]	[redacted]
Age specific manifestation rate of Advanced lung disease	[redacted]	[redacted]	[redacted]
Long term QoL impact of lymphoproliferation + splenomegaly for SoC	[redacted]	[redacted]	[redacted]
Leniolisib costs	[redacted]	[redacted]	[redacted]
Bronchiectasis associated airway disease utility multiplier for leniolisib	[redacted]	[redacted]	[redacted]
Age specific manifestation rate of Cytopenia	[redacted]	[redacted]	[redacted]

HR of Immunoglobulin replacement therapy (IGRT) yr 5+	[redacted]	[redacted]	[redacted]
Age specific manifestation rate of Lymphoproliferation	[redacted]	[redacted]	[redacted]
Bronchiectasis associated airway disease utility multiplier for SoC	[redacted]	[redacted]	[redacted]
Resolution of manifestation for Lymphoproliferation	[redacted]	[redacted]	[redacted]

Source: CS Doc B, Section 3.10 [14]
Abbreviations: CS = company submission; ICER = Incremental Cost-Effectiveness Ratio; HR = Hazard Ratio; OWSA = one-way sensitivity analysis; QoL = Quality of Life; SoC = Standard of Care;

Figure 2: Results of the OWSA

[redacted]

Source: CS Doc B, Section 3.10 [14]
Abbreviations: HR: hazard ratio, ICER: incremental cost-effectiveness ratio, IGRT: immunoglobulin replacement therapy (IRT), SoC: standard of care (current clinical management).

6.2.3 Scenario analysis

Scenario results conducted by the company are summarised in Table 32.

The scenario analyses conducted by the company suggest:

- Using the modified SEE clinician estimates had the highest impact on the ICER. This reduced the cost-effectiveness of leniolisib and increased the ICER by 50%, to [redacted]
- Removal of the age-related utility decrements applied within the base case to reflect a gradual decline in HRQoL with age, as seen in the general population, resulted in the biggest improvement in cost-effectiveness, reducing the ICER to [redacted].

Table 32: Results of deterministic scenario analysis results for the company base-case (with QALY weighting and proposed PAS)

#	Model aspect	Base-case	Scenario analysis	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
	Base case			[redacted]	10.46	[redacted]
1	Source of overall morality for the current clinical management	Case series identified by an SLR	Manifestation-specific mortality	[redacted]	10.18	[redacted]
2	Source of manifestation rates under current clinical management	the cohort of individuals with APDS in the ESID registry data [36] and trial data (Study 2201 Part II) [22]	modified SEE clinician estimate [39]	[redacted]	10.26	[redacted]
3	Impact of leniolisib on manifestations	Various evidence sources, including: the leniolisib trials (Study 2201 Part II and 2201E1 [22 25], the lleniolisib EAP,[31] and a modified SEE clinician estimate [39]	modified SEE clinician estimate [39]	[redacted]	8.51	[redacted]
4	Resource use reduction for manifestations with reduced severity	50%	25%	[redacted]	10.46	[redacted]
5	Age-related utility decrements	Yes	No	[redacted]	11.01	[redacted]
6	Source of utility data for manifestations	SLR	clinician EQ-5D vignette study [45]	[redacted]	9.95	[redacted]

7	Source of baseline utility	EQ-5D vignette valuation exercise	Study 2201 Part II [22] (SF-36 mapped to EQ-5D-3L)	[redacted]	10.24	[redacted]
8	Source of baseline utility	EQ-5D vignette valuation exercise	general population estimate by Ara&Brazier (2010) [46]	[redacted]	10.52	[redacted]
9	Utility impact reduction for manifestations with reduced severity	50%	25%	[redacted]	9.97	[redacted]
10	Treatment discontinuation rate	Study 2201, Study 2201E1 and the leniolisib EAP [31]	Clinician estimate [39]	[redacted]	5.14	[redacted]

Source: CS Doc B, Section 3.10 [14]
Abbreviations: AE = adverse event; APDS = Activated PI3K delta syndrome; ESID = European Society for Immunodeficiencies; ICER; = incremental cost-effectiveness ratio; QALY = Quality Adjusted Life Years; SLR = Systematic Literature Review

6.3 Model validation and face validity check

6.3.1 Face validity assessment and technical verification

The model has gone through a technical verification process by two separate and independent health economist experts.

6.3.2 Comparison with external data

No external data was used to validate the outcomes from the model.

7 EVIDENCE ASSESSMENT GROUP'S ADDITIONAL ANALYSES

7.1 Exploratory and sensitivity analyses undertaken by the EAG

Based on the considerations discussed in the preceding sections above, the EAG base-case included several adjustments to the company base-case presented in Section 6. These adjustments have been subdivided into three categories (derived from Kaltenthaler 2016).

- Fixing errors (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (correcting the model where the EAG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (amending the model where the EAG considers that reasonable alternative assumptions are preferred)

7.1.1 EAG base-case

Adjustments made to derive the EAG base-case (using the CS base-case as starting point) are listed below.

Fixing errors

- The pack cost for leniolisib

Excel file: the company stated that pack cost for leniolisib is [redacted], but this is inconsistent with the information in Table 2, Section B.1.2 in the CS which states that "The anticipated list price of leniolisib is [redacted] per pack of 60 tablets, excluding VAT." The EAG have communicated with NICE and confirmed that the correct price per pack should be [redacted] as per their internal records. Therefore, the value in "Cost" Sheet, Cell J16, changed to [redacted].

- The EAG were not able to verify some of the unit costs submitted by the company. An updated list of the unit costs applied to the economic model has been included in Appendix 1.

- Excel file: some of the costs the company used are for the under-18 only but applied to patients of all ages (e.g., Endoscopy under the Gastrointestinal disorders manifestation), The EAG has modified the formula so that it accurately captured the different unit costs applied for resource use applicable to patients under 18 years old and over 19 years old. The EAG has added 18-over costs for these treatments in the "parameter" sheet, and applied the under-18 and over-19 costs where appropriate in the leniolisib and standard of care model engines.

Fixing violations

As mentioned in Section 5.2.2.4, QALY gain weight should not be applied in the base-case as the decision of whether the submission meets the criteria for the QALY gain weight and the magnitude of the weight to be applied (and consequently what ICER threshold to use) should be made by the NICE committee. Therefore, the EAG has presented the unweighted results in the EAG base-case analysis but present the weighted results in a scenario analysis if the results suggest that a QALY gain weight can be applied. As a result, the EAG have modified the following values in the submitted excel economic model: "Results" Sheet, Cell G23 and J23, were changed from 1.5 to 1 for the base-case analysis.

Matters of judgement

- The assumption of the discount rate applied to both costs and health effects

As detailed in Section 5.2.2.3, The EAG acknowledge the potentially large positive impact of leniolisib on the improvement of QoL and life expectancy of people with APDS, yet also uncertainty in the long term effectiveness of leniolisib remains. Therefore, the EAG applied a 3.5% discount rate for both the costs and health effects in the EAG base-case analysis, and conducted a scenario using a 1.5% discount rate for both the costs and health effects in the EAG's scenario analysis. This means in the Excel file, the value in "Setting" Sheet, Cell K26, was changed from 1.5% to 3.5% for the base-case.

- Treatment discontinuation rate

As discussed in Section 5.2.4.1, the EAG think that the treatment discontinuation rate used in the CS's base case is too low, and thus adopted a higher discontinuation rate of14% (point estimate of expert elicitation exercise).

- The additional utility gain assumed for the psychological impact of leniolisib.

As discussed in Section 5.2.6.3, the EAG believe that the evidence used to justify the utility gain due to the psychological impact of leniolisib is insufficient and associated with a high degree of uncertainty, which may bias the CE results. Therefore, the EAG have removed this assumption from the EAG base-case analysis, which means the value in "Utility" Sheet, Cell F17, changed from 0.1 to 0.

- The assumption of using a 20% of the mean estimates as standard errors for input parameters where no information on uncertainty is available.

Most of the input parameters in the CS have no information on uncertainty and thus the company made a 10% SE assumption on these parameters in the probabilistic analysis. However, the EAG considered the justification for using a 10% SE insufficient, and therefore adopted a 20% SE in the EAG probabilistic analysis, which is more conservative than the one used in the CS (see Section 5.2.3.2 for further discussion).

7.1.2 EAG exploratory scenario analyses

The EAG performed the following scenario analyses to explore the impact of alternative assumptions conditional on the EAG base-case.

EAG scenarios

(1) Assuming the treatment discontinuation rate to be 0%, as opposed to 14% in the EAG base case (see section 5.2.4.1)
(2) Assuming the treatment discontinuation rate to be 10%, as opposed to 14% in the EAG base case (see section 5.2.4.1)
(3) Assuming the treatment discontinuation rate to be 30%, as opposed to 14% in the EAG base case (see section 5.2.4.1)
(4) Assuming a 1.5% discount rate for both the cost and health effects of the model, as opposed to 3.5% in the EAG base case (see section 5.2.2.3)
(5) Assuming a starting age of 18 rather than 15 (see section 5.2.2.1)
(6) Assuming a 10% standard error (same as CS) for input parameters without information on uncertainty (probabilistic analysis only)

Scenarios from the CS

(7) Assuming no age-related utility decrement.
(8) Utility of manifestations: using utility values associated with each manifestation from the clinician EQ-5D exercise (see section 5.2.6.2)
(9) Utility of manifestations: assuming the utility impact reduction for manifestations with reduced severity being 25% (see section 5.2.6.2)
(10) Baseline utility: using the baseline utility value informed by data from Study 2201 Part II. [22] Baseline SF-36 data from Study 2201 Part II were mapped to EQ-5D-3L utilities using the mapping algorithm reported by Brazier & Rowen (2009).[7] This leads to a baseline utility of [redacted] (SE assumed to be [redacted]) (see section 5.2.6.2)
(11) Baseline utility: general population utility values are calculated for each cohort using methods set out by Ara & Brazier (2010)[46]. This leads to a baseline utility of [redacted] (SE: [redacted]) (see section 5.2.6.2)
(12) Utility of manifestations: Using the lower bound of the utility value elicited from the Pharming TTO as the lower limit on utilities for this model (see section 5.2.6.2)

7.1.3 EAG subgroup analyses

No additional subgroup analyses were conducted by the EAG.

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

7.2.1 The EAG base case analysis

Table 33 reports the cost-effectiveness results of updating the company base-case model correcting for errors found by the EAG, correcting violations identified by EAG and the individual impact of the matters of judgement by the EAG to generate the EAG base-case results. Once errors and violations have been corrected in the company's base-case model the unweighted deterministic ICER increases from [redacted] to [redacted]. The unweighted probabilistic ICER resulting from the company's base-case analysis slightly decreases from [redacted] to [redacted]

After fixing errors and correcting violations in the company model the impact of the EAG preferred assumptions applied to the company's model is also detailed in Table 33 and summarised below:

- Applying a 3.5% rather than a 1.5% discount rate for the health effects in the model has the biggest impact on the cost-effectiveness results, increasing the ICER amount by about [redacted] (to [redacted]). Compared with the CS base, QALYs for both the leniolisib and standard of care (SoC) arm decreased as did the difference in QALYs between the two arms.
- The removal of the utility gain associated with the positive psychological impact of initiating leniolisib treatment has the next biggest impact on cost-effectiveness resulting in an ICER of [redacted].
- Applying a larger treatment discontinuation rate (mean value derived from the company's expert opinion exercise) greatly decreases both the cost and QALYs for the leniolisib arm, resulting in a drop in the ICER to [redacted].
- Assuming a 20% rather than a 10% SE for input parameters without information on uncertainty available has little impact on the probabilistic cost-effectiveness results

The undiscounted QALY gain from the preferred EAG base-case analysis is 5.86. As this QALY gain is less than 10, the EAG note that the NICE criteria for the application of a QALY gain weight is not met on this occasion.

Table 33: Deterministic/probabilistic EAG base-case

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS base-case – deterministic (with QALY gain weight)					
Leniolisib	[redacted]	[redacted]	[redacted]	15.46 (weighted)	[redacted]
SoC	1,587,334	[redacted]			
CS base-case – probabilistic (with QALY gain weight)					
Leniolisib	[redacted]	[redacted]	[redacted]	17.10 (weighted)	[redacted]
SoC	1,613,679	[redacted]			
CS base-case – deterministic (without QALY gain weight)					
Leniolisib	[redacted]	[redacted]	[redacted]	10.46 (unweighted)	[redacted]
SoC	1,587,334	[redacted]			
CS base-case – probabilistic (without QALY gain weight)					
Leniolisib	[redacted]	[redacted]	[redacted]	11.57 (unweighted)	[redacted]
SoC	1,613,679	[redacted]			
Fixing errors (1-3) – deterministic[a]					
Leniolisib	[redacted]	[redacted]	[redacted]	10.46	[redacted]
SoC	1,547,870	[redacted]			
Fixing errors (1-3) – probabilistic[a]					
Leniolisib	[redacted]	[redacted]	[redacted]	11.49	[redacted]
SoC	1,620,167	[redacted]			
Fixing errors (1-3) + fixing violation + applying a 20% standard error for parameters without information on uncertainty – probabilistic[a]					
Leniolisib	[redacted]	[redacted]	[redacted]	11.02	[redacted]
SoC	[redacted]	[redacted]			
Fixing errors (1-3) + fixing violation + applying a 3.5% discount rate to the health effects[a]					
Leniolisib	[redacted]	[redacted]	[redacted]	7.21	[redacted]
SoC	1,547,870	[redacted]			
Fixing errors (1-3) + fixing violation + applying an alternative treatment discontinuation rate elicited from the expert opinion[a]					
Leniolisib	[redacted]	[redacted]	[redacted]	5.14	[redacted]
SoC	1,547,870	[redacted]			
Fixing errors (1-3) + fixing violation + removing the utility gain assumed for the psychological impact of leniolisib[a]					
Leniolisib	[redacted]	[redacted]	[redacted]	8.94	[redacted]
SoC	1,547,870	[redacted]			
EAG base-case (errors 1-3, violation, and matters of judgment 1-3) – deterministic[a]					
Leniolisib	[redacted]	[redacted]	[redacted]	3.54	[redacted]

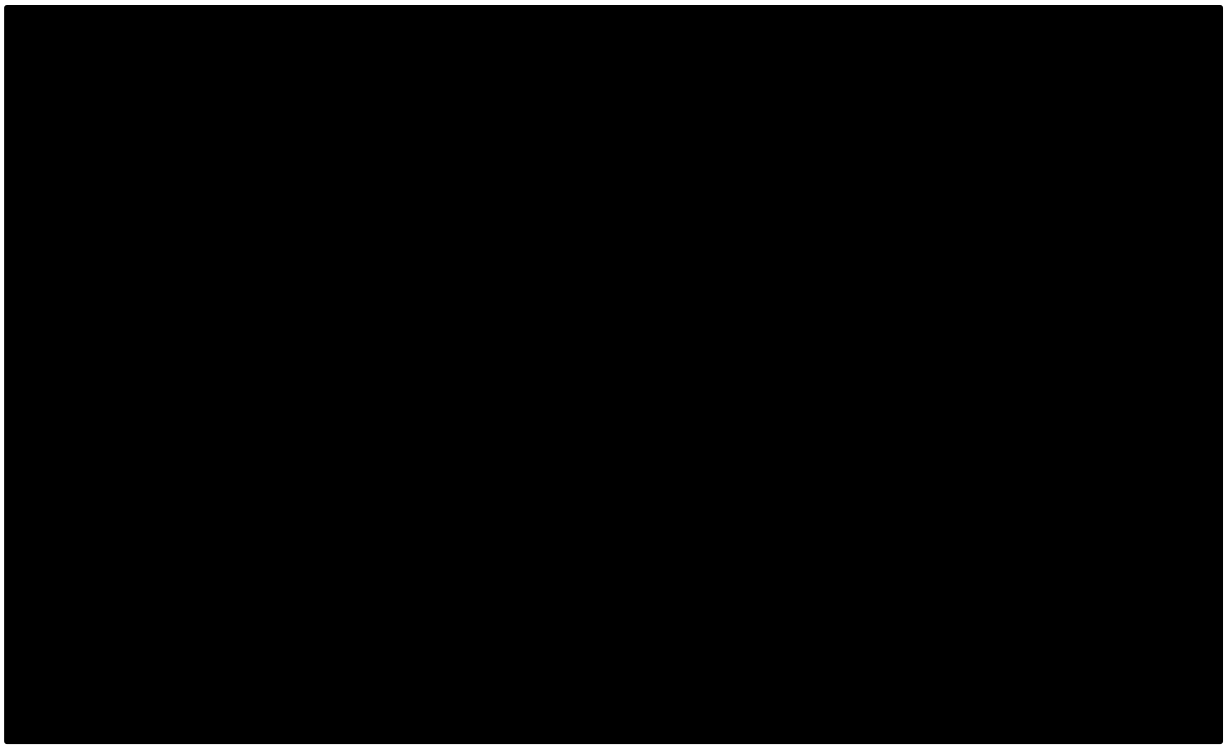
Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
SoC	1,547,870	[redacted]			
EAG base-case (errors 1-3, violation, and matters of judgment 1-3) – probabilistic[a]					
Leniolisib	[redacted]	[redacted]	[redacted]	4.51	[redacted]
SoC	1,646,253	[redacted]			

Abbreviations: CS = company submission; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life year; SoC = Standard of Care
Footnote: (a) Results without the QALY gain weight, as detailed in the section 7.1.1.

The EAG base-case model produced point estimates with accompanying 95% credible intervals in a probabilistic analysis (with 1,000 replications).The estimated EAG base-case ICER), based on the EAG preferred assumptions highlighted in Section 7.1.1, was [redacted] per QALY gained. Incremental QALYs for leniolisib versus current clinical management were 4.51 (95% CrI: -1.76 to 11.78) and incremental costs were [redacted]. The probabilistic EAG base-case analyses suggests that leniolisib has a [redacted] probability of being cost effectiveness at willingness to pay threshold of £100,000 per QALY gained. Therefore, usual care would be favoured in the probabilistic results.

These probabilistic results are shown in the form of a cost-effectiveness plane (Figure 3) and a cost-effectiveness acceptability curve (CEAC) (Figure 4).

Figure 3 Incremental cost-effectiveness plane leniolisib versus current standard of care (EAG base-case)



Source: CS model, EAG's base-case

Abbreviations: GBP = pounds sterling; QALY = quality-adjusted life year

Figure 4 Cost-effectiveness acceptability curve (CEACs) leniolisib versus standard of care (EAG base-case)

Source: CS model, EAG's base-case

Abbreviations: CEAC = cost-effectiveness acceptability curve; CS = company submission; EAG = Evidence Assessment Group; GBP = pounds sterling

7.2.2 The EAG scenario and one-way sensitivity analysis.

Table 34 reports the cost-effectiveness results for the EAG's and CS's scenario analyses. Compared with the EAG's base-case deterministic results. Seven of the scenarios resulted in a lower ICER compared with the EAG base case and six of the scenarios yielded a higher ICER compared with the base case. The scenarios with the largest impact on the cost-effectiveness results assessed by the EAG were:

- Changing the treatment discontinuation rate (SC1 and SC3): Increasing the treatment discontinuation rate from 14% to 30% decreases the incremental cost and decreases the incremental QALYs, resulting in a lower ICER [redacted] vs [redacted]). Decreasing the treatment discontinuation rate from 14% to 0% increases the incremental cost and increases the incremental QALYs, resulting in a higher ICER (£[redacted] vs [redacted]).
- Setting the discount rate to 1.5% for both costs and effect (SC4): Decreasing the discount rate increases the incremental cost slightly and increases the incremental QALYs, resulting in a lower ICEDR (£[redacted] vs [redacted]).
- Setting lower limit on utilities (SC12): This assumption decreases the incremental QALYs significantly, resulting in an increased ICER (£[redacted] vs £[redacted]).

Table 34: EAG scenario analysis results table

Scenario #	EAG base-case input	Alternative input	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
	EAG base-case (deterministic)	N/A	[redacted]	3.54	[redacted]
	EAG base-case (probabilistic)	N/A	[redacted]	4.51	[redacted]
1	Treatment discontinuation rate = 14%	Treatment discontinuation rate = 0%	[redacted]	9.49	[redacted]
2		Treatment discontinuation rate = 10%	[redacted]	4.04	[redacted]
3		Treatment discontinuation rate = 30%	[redacted]	2.80	[redacted]
4	Discount rate = 3.5% for both costs and health effects	Discount rate =1.5% for both costs and health effects	[redacted]	4.62	[redacted]
5	Starting age = 15	Starting age = 18	[redacted]	3.72	[redacted]
6 [a]	A 20% standard error assumption on parameters without information on uncertainty	A 10% standard error assumption on parameters without information on uncertainty	[redacted]	4.49	[redacted]
7	Assuming age-related utility decrement	Assuming no age-related utility decrement	[redacted]	3.61	[redacted]
8	Using utility values for each manifestation from the literature	Using utility values for each manifestation from the clinician EQ-5D exercise	[redacted]	3.65	[redacted]
9	Assuming the utility impact reduction for manifestations with reduced severity being 50%	Assuming the utility impact reduction for manifestations with reduced severity being 25%	[redacted]	3.40	[redacted]
10	Baseline utility informed by the clinician's estimates	Baseline utility informed by the trial data	[redacted]	3.45	[redacted]
11		Baseline utility informed by	[redacted]	3.79	[redacted]

Scenario #	EAG base-case input	Alternative input	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
		general population utility values by Ara & Brazier (2010) [46].			
12	No lower limit on utilities	Lower limit on utilities elicited from TTO tasks	[redacted]	3.12	[redacted]

Source: EAG outputs
Abbreviations: EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; N/A = not applicable; QALY = quality-adjusted life year; TTO = Time trade-off
Table note: (a) Probabilistic analysis results are reported

The EAG conducted one-way sensitivity analyses (OWSA) by varying a selection of model parameters individually. As mentioned in section 5.2.3.2, The EAG prefers a more conservative value of the SE assumption for sensitivity analysis. Therefore, for parameters where empirical 95% CIs were not available, a SE of 20% of the mean was assumed in the EAG OWSA. Parameters were varied within lower and upper bounds set to 2.5% and 97.5% of their 95% CIs.

The results of the OWSA are presented in Table 35 and Figure 5. The parameters with the greatest influence on the ICER are leniolisib costs, the HR of Immunoglobulin replacement therapy (IGRT) yr 5+ and the age specific manifestation rate of Gastrointestinal manifestations. Nine of the top ten parameters included in Table 35 are the same as those included in the company's own top ten most sensitive parameters (see Table 31). The only difference is that the EAG list now includes the "subsequent years discontinuation rate" parameter as having a big influence on the ICER. This is to be expected as the EAG preferred analysis base-case analysis included a 14% discontinuation rate which is considerably higher than the 3.54% discontinuation rate included by the company in their economic model.

Table 35 - OWSA results for leniolisib versus current clinical management

Parameter name	Lower bound ICER (£)	Upper bound ICER (£)	Difference (£)
Leniolisib costs	[redacted]	[redacted]	[redacted]
HR of Immunoglobulin replacement therapy (IGRT) yr 5+	[redacted]	[redacted]	[redacted]
Age specific manifestation rate of Gastrointestinal manifestations	[redacted]	[redacted]	[redacted]
Long term QoL impact of lymphoproliferation + splenomegaly for SoC	[redacted]	[redacted]	[redacted]
Age specific manifestation rate of Advanced lung disease	[redacted]	[redacted]	[redacted]
Age specific manifestation rate of Cytopenia	[redacted]	[redacted]	[redacted]
Subsequent years discontinuation rate	[redacted]	[redacted]	[redacted]

Age specific manifestation rate of Lymphoproliferation	[redacted]	[redacted]	[redacted]
Bronchiectasis associated airway disease utility multiplier for leniolisib	[redacted]	[redacted]	[redacted]
Gastrointestinal disorders (GI) utility multiplier for SoC	[redacted]	[redacted]	[redacted]

Abbreviations: CS = company submission; ICER = Incremental Cost-Effectiveness Ratio; HR = Hazard Ratio; OWSA = one-way sensitivity analysis; QoL = Quality of Life; SoC = Standard of Care;

Figure 5 Results of the EAG OWSA

[redacted]

7.3 Overall conclusions of the EAG's cost-effectiveness analysis

The EAG base-case fixed the errors in the pack cost for leniolisib, updated several other unit costs applied to the model, and changed the approach the company calculated the age-dependent costs. The EAG base-case also changed the 1.5% discount rate assumption to 3.5% based on the NICE reference case.[1] Other preferred assumptions incorporates into the EAG base case include using an alternative source for the treatment discontinuation rate and removing the additional utility gain due to psychological effects of leniolisib on quality of life.

The EAG base case (probabilistic) results comparing leniolisib with current clinical management yielded 4.51(95% CrI: -1.76 to 11.78) incremental QALYs [redacted] incremental costs. This resulted in an ICER of [redacted] per QALY gained. The probabilistic EAG base-case analyses suggests that leniolisib has a [redacted] probability of being cost effectiveness at willingness to pay threshold of £100,000 per QALY gained. However, the wide confidence intervals suggest a high degree of uncertainty surrounding costs and effects.

The parameters with the greatest influence on the ICER are leniolisib costs, the HR of Immunoglobulin replacement therapy (IGRT) after Year 5 and the age-specific manifestation rate of Gastrointestinal manifestations were found by the EAG to be the parameters with the largest impact on the cost-effectiveness results in the one-way sensitivity analysis.

Treatment discontinuation has the biggest impact on the cost-effectiveness results. For example, increasing the treatment discontinuation rate from 14% to 30% decreases leniolisib treatment costs, significantly reducing the incremental costs from [redacted] to [redacted] whilst also decreasing the incremental QALYs from 4.51 to 2.80 the incremental QALYs, resulting in an ICER of [redacted].

7.4 Overall conclusions of the EAG's critique

7.4.1 Clinical effectiveness

The SLR to identify all relevant clinical evidence on the safety and efficacy of leniolisib to treat patients with APDS was last updated with a targeted update in April 2024. It identified 30 observational or interventional studies and 88 case studies, and the EAG believes it has captured all related evidence relating to the decision problem.

Study 2201 part I (dose-finding study, n=6) provided no major concerns for the EAG, and concluded that 70mg bid was the appropriate dose for the population under consideration.

Study 2201 Part II (RCT, n=31) appears to have been methodologically sound although some areas, such as concealment of allocation, are at unclear risk of bias. The key issue with the RCT is that the comparator group did not receive established clinical management as understood in the UK and defined in the NICE scope. They received a placebo plus restricted symptomatic management but selected immunosuppressants (which the EAG's clinical experts considered reflect standard care) were prohibited, which may have over-estimated the apparent effectiveness of leniolisib. The small sample size (n=31) is appropriate relative to the estimated number of people living with APDS and the rarity of the condition. Part II reported a statistically significant and clinically meaningful change in surrogate co-primary endpoints used to measure immunophenotype normalisation (increased percentage naïve B cells of total B cells) and reduction in lymphoproliferation (change in size of lesions) respectively.

Study 2201E1 (open-label extension trial, n=37) provided leniolisib to participants from part I and part II, plus two other eligible participants. It is due to complete in 2027 and has released interim results, reporting continuing immunophenotype normalisation to day 252, reduction in serum IgM levels, (post-hoc) reduction in the incidence of infections, decreased fatigue, and improved within-patient quality of life scores. It also reported that leniolisib continued to be well tolerated throughout a median of 154.71 weeks.

To provide further evidence the company carried out an indirect treatment comparison which compared leniolisib patients from the extension trial to a real-world sample of control patients from the ESID Registry, who did not have their treatments prohibited (as in the trial). This reported reductions in infection rate and reduction (improvement) in serum IgM levels.

7.4.2 Cost effectiveness

The EAG considers that the company's deviations from the reference case had a large impact on cost-effectiveness results. This is most evident when the discount rate of 3.5% was applied to both costs and effects as per the NICE reference case.

Lack of long-term efficacy and quality of life data was a concern. The EAG appreciates the company was hampered by the lack of data on long term efficacy and quality of life data and sought alternative sources, including a very thorough expert elicitation exercise, to seek the data needed for the economic model. For example, the company had to rely on proxy conditions to apply their manifestation-related utility and expert opinion as there were no useable HRQoL data that could be directly incorporated into the economic model. This is unsurprising given the rarity of the condition and the small number of patients affected by APDS in the UK. This led to a number of assumptions which incorporated a high degree of

uncertainty into the analyses. The EAG are sympathetic to this approach and recognises that emerging longer-term data will be needed to address this uncertainty.

The most influential cost driver in the EAG analysis was the cost of leniolisib itself. This cost is based upon the confidential PAS cost. All analyses used this cost. The costs for each of the health states in the model related to the care of participants from the age of 15. The EAG had some concerns surrounding resource use. Most of the estimates relating to healthcare resource use were derived via an expert elicitation exercise. The EAG note that the wide variation in the estimates provided by the experts leads to uncertainty surrounding the management and monitoring of APDS for both leniolisib and current clinical management patients. Additionally, the analysis was restricted to the NHS and did not include any resources associated with the use of personal and social care services. This was a concern as the manifestations associated with APDS can severely affect the patient's daily activities such as education and work. The EAG note that APDS patients may therefore need extra support that could potentially be provided by personal and social care services and which may increase the costs associated with the condition.

Whilst the probabilistic EAG base-case analyses suggests that leniolisib has a [redacted] probability of being cost effective at willingness to pay threshold of £100,000 per QALY gained, these results carry a high degree of uncertainty surrounding costs and effects suggesting that more research is needed. The EAG analyses also show that some changes in the assumptions incorporated in the model have a substantial impact on the relative cost-effectiveness of leniolisib.

8 REFERENCES

1 NICE. *NICE health technology evaluations: the manual. Process and methods [PMG36]. Last Update Date: 31 January 2022*. London: National Institute for Health and Care Excellence (NICE); 2022. Available from: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation [Accessed 28/05/2024].

2 NICE. *NICE health technology evaluations: the manual. Process and methods [PMG24].* London: National Institute for Health and Care Excellence (NICE); 2022. Available from: https://www.nice.org.uk/process/pmg24 [Accessed 9th July 2024].

3 Lanitis T, Muszbek N, Tichy E. *The Probability of a Successful Probabilistic Sensitivity Analysis* London: Evidera; 2014. Available from: https://www.evidera.com/wp-content/uploads/2015/04/The-Probability-of-a-Successful-Probabilistic-Sensitivity-Analysis.pdf [Accessed 10th July 2024].

4 Novartis Pharmaceuticals. *An Open-label, Non-randomized, Within-patient Dose-finding Study Followed by a Randomized, Subject, Investigator and Sponsor Blinded Placebo Controlled Study to Assess the Efficacy and Safety of CDZ173 (Leniolisib) in Patients With APDS/PASLI (Activated Phosphoinositide 3-kinase Delta Syndrome/ p110δ-activating Mutation Causing Senescent T Cells, Lymphadenopathy and Immunodeficiency)*. 2015. Available from: https://clinicaltrials.gov/study/NCT02435173 (login required) [Accessed 19 July 2024].

5 Alava M, Pudney S, Wailoo A. *Estimating the relationship between EQ-5D-5L and EQ-5D-3L: results from an english population study*. Sheffield: Health Economics and Decision Science, School of Health and Related Research. University of Sheffield; 2020. Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/estimating-the-relationship-betweenE-Q-5D-5L-and-EQ-5D-3L.pdf [Accessed 10th July 2024].

6 Pharming Technologies B.V. *Leniolisib for activated phosphoinositide 3-kinase delta syndrome in people 12 years and over [ID6130]. Cost-effectiveness model. Highly specialised technology appraisal*. 2024. Available from: [Accessed 10th July 2024].

7 Brazier J, Rowen D. NICE Decision Support Unit Technical Support Documents. In: *NICE DSU Technical Support Document 11: Alternatives to EQ-5D for Generating Health State Utility Values*. London: National Institute for Health and Care Excellence (NICE); 2011.

8 Rao VK, Webster S, Dalm VASH, Šedivá A, Van Hagen PM, Holland S, et al. Effective “activated PI3Kδ syndrome”–targeted therapy with the PI3Kδ inhibitor leniolisib. *Blood*. 2017;130(21):2307-16. Available from: https://doi.org/10.1182/blood-2017-08-801191.

9 Maccari ME, Abolhassani H, Aghamohammadi A, Aiuti A, Aleinikova O, Bangs C, et al. Disease Evolution and Response to Rapamycin in Activated Phosphoinositide 3-Kinase δ Syndrome: The European Society for Immunodeficiencies-Activated Phosphoinositide 3-Kinase δ Syndrome Registry. *Frontiers in Immunology*. 2018;9. Available from: https://doi.org/10.3389/fimmu.2018.00543.

10 Singh A, Joshi V, Jindal AK, Mathew B, Rawat A. An updated review on activated PI3 kinase delta syndrome (APDS). *Genes & Diseases*. 2020;7(1):67-74. Available from: https://doi.org/https://doi.org/10.1016/j.gendis.2019.09.015.

11 Coulter TI, Cant AJ. The Treatment of Activated PI3Kδ Syndrome. *Frontiers in Immunology*. 2018;9. Available from: https://doi.org/10.3389/fimmu.2018.02043.

12 Ayuk F, Balduzzi A. Donor Selection for Adults and Pediatrics. In. Springer International Publishing; 2019: 87-97.

13 NHS. *Children and young people. Consent to treatment.* 2022. Available from: https://www.nhs.uk/conditions/consent-to-treatment/children/.

14 Pharming Technologies B.V. *Leniolisib for activated phosphoinositide 3-kinase delta syndrome (APDS) in people 12 years and over [ID6130] Document B. Company Evidence Submission. Highly specialised technology appraisal*. Pharming Technologies BV; 2024.

15 NICE. *Approximate Conversions and Units*. 2024. Available from: https://bnf.nice.org.uk/about/approximate-conversions-and-units/ [Accessed 9th July 2024].

16 Pharming Technologies B.V. *Pediatric Patients Aged 4 to 11 Years With APDS. NCT05438407*. 2023. Available from: https://clinicaltrials.gov/study/NCT05438407 [Accessed 17 July 2024].

17 Pharming Technologies B.V. *An Open-label, Single Arm Study of the Safety, Pharmacokinetics, Pharmacodynamics, and Efficacy of Leniolisib in Pediatric Patients (Aged 1 to 6 Years) With APDS (Activated Phosphoinositide 3-Kinase Delta Syndrome) Followed by an Open-label Long-term Extension*. 2022. Available from: https://clinicaltrials.gov/study/NCT05693129 [Accessed 17 July 2024] (login required) [Accessed 2023/08/30/].

18 McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *Journal of Clinical Epidemiology*. 2016;75:40-6. Available from: https://doi.org/10.1016/j.jclinepi.2016.01.021.

19 Pharming Technologies B.V. *Leniolisib for activated phosphoinositide 3-kinase delta syndrome (APDS) in people 12 years and over [ID6130] Appendices. Company Evidence Submission. Highly specialised technology appraisal*. Pharming Technologies BV; 2024.

20 Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health*. 1998;52(6):377-84. Available from: https://doi.org/10.1136/jech.52.6.377.

21 Critical Appraisal Skills Programme (CASP). *CASP cohort study checklist*. Oxford, UK; 2018. Available from: https://casp-uk.net/checklists/casp-cohort-studies-checklist-fillable.pdf [Accessed 10th July 2024].

22 Rao VK, Webster S, Šedivá A, Plebani A, Schuetz C, Shcherbina A, et al. A randomized, placebo-controlled phase 3 trial of the PI3Kδ inhibitor leniolisib for activated PI3Kδ syndrome. *Blood*. 2023;141(9):971-83. Available from: https://doi.org/10.1182/blood.2022018546.

23 Novartis. *An open-label, non-randomized, within-patient dose-finding study followed by a randomized, double-blind placebo controlled study with extension to assess the safety and efficacy of CDZ173 in patients with APDS/PASLI (Activated phosphoinositide 3-kinase delta syndrome/ p110δ- activating mutation causing senescent T cells, lymphadenopathy and immunodeficiency). Clinical study report part II*. Novartis Institutes for Biomedical Research; 2020. Available from: [Accessed 11th July 2024].

24 Christ A, Laurence C, De Buck S, Aniket J, Pylvaenaeinen I, Ulrike S, et al. *An open-label, non-randomized, within-patient dose-finding study followed by a randomized, double-blind placebo controlled study with extension to assess the safety and efficacy of CDZ173 in patients with APDS/PASLI (Activated phosphoinositide 3-kinase delta syndrome/ p110δ- activating mutation causing senescent T cells, lymphadenopathy and immunodeficiency). Protocol*. Novartis Institutes for Biomedical Research; 2014. Available from: [Accessed 10th July 2024].

25 Rao VK, Kulm E, Šedivá A, Plebani A, Schuetz C, Shcherbina A, et al. Interim analysis: Open-label extension study of leniolisib for patients with APDS. *Journal of Allergy and Clinical Immunology*. 2023;153(1):265-74.e9. Available from: https://doi.org/10.1016/j.jaci.2023.09.032.

26 Pharming Technologies B.V. *Leniolisib for activated phosphoinositide 3-kinase delta syndrome in people 12 years and over [ID6130]. Points for clarification response. Highly specialised technology appraisal*. Pharming Technologies BV,; 2024.

27 Youssef J, Novosad SA, Winthrop KL. Infection Risk and Safety of Corticosteroid Use. *Rheumatic Disease Clinics of North America*. 2016;42(1):157-76. Available from: https://doi.org/10.1016/j.rdc.2015.08.004.

28 van Gent R, van Tilburg CM, Nibbelke EE, Otto SA, Gaiser JF, Janssens-Korpela PL, et al. Refined characterization and reference values of the pediatric T- and B-cell compartments. *Clinical Immunology*. 2009;133(1):95-107. Available from: https://doi.org/10.1016/j.clim.2009.05.020.

29 Morbach H, Eichhorn EM, Liese JG, Girschick HJ. Reference values for B cell subpopulations from infancy to adulthood. *Clinical and Experimental Immunology*. 2010;162(2):271-9. Available from: https://doi.org/10.1111/j.1365-2249.2010.04206.x.

30 Rao VK, Kulm E, Grossman J, Buchbinder D, Chong H, Bradt J, et al. Long-term treatment with selective PI3Kδ inhibitor leniolisib in adults with activated PI3Kδ syndrome. *Blood Advances*. 2024;8(12):3092-108. Available from: https://doi.org/10.1182/bloodadvances.2023011000.

31 Pharming Group. *A Survey of Physicians Treating APDS Patients with Leniolisib Through Early Access Supply/Program: A Semi-Quantitative Evaluation of Response to Treatment. [Accessed 10th July 2024]*. 2023.

32 Whalen J, Chandra A, Kracker S, Ehl S, Seidel MG, Gulas I, et al. Comparative efficacy of leniolisib (CDZ173) versus standard of care on rates of respiratory tract infection and serum immunoglobulin M (IgM) levels among individuals with activated phosphoinositide 3-kinase delta (PI3Kδ) syndrome (APDS): An externally-controlled study. [Manuscript]. 2024. [Accessed 9th July 2024].

33 NHS England. *2021/22 National Cost Collection Data Publication*. 2023. Available from: https://www.england.nhs.uk/wp-content/uploads/2023/04/3_The_national_cost_collection_index_NCCI_FY21-22_v4-2.xlsx [Accessed 9th July 2024].

34 NICE. *British National Formulary (BNF)*. 2024. Available from: https://bnf.nice.org.uk/ [Accessed 9th July 2024].

35 Department of Health and Social Care. *Guidance: Drugs and pharmaceutical electronic market information tool (eMIT)*. 2011. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit [Accessed 9th July 2024].

36 Costello Medical. *ESID registry analysis. APDS patient characteristics. Analysis summary document*. 2024. Available from: [Accessed 10th May 2024].

37 NICE. *Sebelipase alfa for treating Wolman disease. Highly specialised technologies guidance.* National Institute for Health and Care Excellence (NICE); 2024. Available from: https://www.nice.org.uk/guidance/hst30/resources/sebelipase-alfa-for-treating-wolman-disease-pdf-50216373062341 [Accessed 24 July 2024].

38 Harrington A, Mahendran M, Ramasubramanian R, Memmott H, Germain G, Büsch K, et al. OVERALL SURVIVAL (OS) AMONG PATIENTS WITH ACTIVATED PHOSPHOINOSITIDE 3-KINASE DELTA SYNDROME (APDS). *Annals of Allergy, Asthma & Immunology*. 2023;131(5):S63-S4. Available from: https://doi.org/10.1016/j.anai.2023.08.190.

39 UKepic. *Leniolisib Expert Perspective and Insight Consultancy to Inform UK Health Technology Assessment. Final Report. [Accessed 10th July 2024]*. 2024.

40 Pharming Technologies B.V. *Periodic safety update report / periodic benefit-risk evaluation report for active substance(s): Leniolisib ATC code(s): L03ax22 (May 2024). [Accessed 10th July 2024]*. 2024.

41 Pharming Technologies B.V. *Periodic safety update report / periodic benefit-risk evaluation report for active substance(s): Leniolisib ATC code(s): L03AX22 (November 2023). [Accessed 10th July 2024]*. 2023.

42 Pharming Technologies B.V. *An open-label, non-randomized extension study to evaluate the long-term safety, tolerability, efficacy, and pharmacokinetics of leniolisib in patients with APDS/PASLI (Activated phosphoinositide 3-kinase delta syndrome/p110δ-activating mutation causing senescent T cells, lymphadenopathy and immunodeficiency) [CCDZ173X2201E1][Accessed 10th July 2024]*. 2023.

43 Horscroft J, Lumanity; Lee D, PenTAG, The University of Exeter; Jankovic D, The University of York; Soares M, The University of York; Bojke L, et al. *Structured expert elicitation for healthcare decision making: A practical guide*. United Kingdom; 2024. Available from: https://www.york.ac.uk/media/che/documents/Structured%20expert%20elicitation%20for%20healthcare%20decision%20making%20A%20practical%20guide.pdf [Accessed 9th July 2024].

44 Odnoletkova I, Kindle G, Quinti I, Grimbacher B, Knerr V, Gathmann B, et al. The burden of common variable immunodeficiency disorders: a retrospective analysis of the European Society for Immunodeficiency (ESID) registry data. *Orphanet Journal of Rare Diseases*. 2018;13(1). Available from: https://doi.org/10.1186/s13023-018-0941-0.

45 Adelphi Values PROVE™. *Healthcare professional (HCP) valuation of health states in APDS using EQ 5D-5L. [Accessed 10th July 2024]*. 2024.

46 Ara R, Brazier J. Deriving an algorithm to convert the eight mean sf-36 dimension scores into a mean eq-5d preference-based score from published studies (where patient level data are not available). *Value in Health*. 2008;11(7):1131-43. Available from: https://doi.org/10.1111/j.1524-4733.2008.00352.x.

47 Kind P, Hardman G, Macran S. *UK population norms for EQ-5D. Discussion paper 172*. York: Centre for Health Economics; 1999. Available from: https://www.york.ac.uk/che/pdf/DP172.pdf [Accessed 10th July 2024].

48 Rand S, Caiels J. *Using proxies to assess quality of life: A review of the issues and challenges*. Canterbury: QORU. Personal Social Services Research Unit; 2015. Available from: https://www.pssru.ac.uk/pub/4980.pdf [Accessed 10th July 2024].

49 Hedman C, Djarv T, Strang P, Lundgren C. Fear of Recurrence and View of Life Affect Health-Related Quality of Life in Patients with Differentiated Thyroid Carcinoma: A Prospective Swedish Population-Based Study. *Thyroid*. 2018;28(12). Available from: https://doi.org/10.1089/thy.2018.0388.

50 Short H, Al Sayah F, Ohinmaa A, Johnson JA. The performance of the EQ-5D-3L in screening for anxiety and depressive symptoms in hospital and community settings. *Health and Quality of Life Outcomes*. 2021;19(1). Available from: https://doi.org/10.1186/s12955-021-01731-x.

51 Pharming Group N.V. *Clinical Expert Response. Dr Chandra's response to draft scope. [Accessed 10th July 2024]*. 2023.

52 Kung S, Rummans TA, Colligan RC, Clark MM, Sloan JA, Novotny PJ, et al. Association of Optimism-Pessimism With Quality of Life in Patients With Head and Neck and Thyroid Cancers. *Mayo Clinic Proceedings*. 2006;81(12):1545-52. Available from: https://doi.org/10.4065/81.12.1545.

53 Dolan P. Modeling Valuations for EuroQol Health States. *Medical Care*. 1997;35(11):1095-108. Available from: https://doi.org/10.1097/00005650-199711000-00002.

54 Shaw J, Johnson J, Coons S. US Valuation of the EQ-5D Health States: Development and Testing of the D1 Valuation Model. *Medical Care*. 2005;43(3):203-20. https://www.jstor.org/stable/3768219 [Accessed 10th July 2024].