Single Technology Appraisal

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Committee Papers

© National Institute for Health and Care Excellence 2026. All rights reserved. See Notice of Rights. The content in this publication is owned by multiple parties and may not be re-used without the permission of the relevant copyright owner.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Contents:

The following documents are made available to stakeholders:

1. **Comments on the Draft Guidance from Merck Sharp & Dohme**
   a. Draft Guidance Comments from Merck Sharp & Dohme
   b. Additional clarification questions from the EAG following the company's Draft Guidance response, and company's response to clarification questions

2. **Comments on the Draft Guidance from experts:**
   a. Dr Daniel Knight, Consultant Cardiologist Lead for the Royal Free National Pulmonary Hypertension Service - Clinical Expert, nominated by Merck Sharp & Dohme

3. **Comments on the Draft Guidance received through the NICE website**

4. **External Assessment Group critique of company response to the DG:**
   a. EAG Critique
   b. Updated severity weighting calculations

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

© National Institute for Health and Care Excellence 2026. All rights reserved. See Notice of Rights. The content in this publication is owned by multiple parties and may not be re-used without the permission of the relevant copyright owner.

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

	Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly. The Appraisal Committee is interested in receiving comments on the following: • has all of the relevant evidence been taken into account? • are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? • are the provisional recommendations sound and a suitable basis for guidance to the NHS? NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the preliminary recommendations may need changing in order to meet these aims. In particular, please tell us if the preliminary recommendations: • could have a different impact on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology; • could have any adverse impact on people with a particular disability or disabilities. Please provide any relevant information or data you have regarding such impacts and how they could be avoided or reduced.
Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):	Merck Sharp & Dohme (UK) Ltd
Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months. [Relevant companies	N/A

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

are listed in the appraisal stakeholder list.] Please state: • the name of the company • the amount • the purpose of funding including whether it related to a product mentioned in the stakeholder list • whether it is ongoing or has ceased.	
Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None
Name of commentator person completing form:	Aidan Byrne; Younan Zhang; Kayla Engelbrecht

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

Comment number	**Comments** Insert each comment in a new row. Do not paste other tables into this table, because your comments could get lost – type directly into this table.

Executive summary	We welcome the opportunity to respond to the draft guidance for sotatercept. Our response contains:
	
	• A response to the committee's requests for further information and analysis which includes:
	  o An updated model structure that allows for improvement on initiation of PCA
	  o Details of a matching adjusted indirect comparison (MAIC) of STELLAR and GRIPHON to further address the committee's concerns regarding the relative effects of sotatercept compared to selexipag
	• A revised company base case that takes into account the committee's preferred assumptions, the revised model structure and the MAIC (plus extensive sensitivity and scenario analysis). Please note guidance for the EAG on the implementation of the new model structure and inputs are provided in Appendix C of this document.
	• Corrections to some statements made in the DG that we understand to be incorrect
	
	We urge the committee to consider our revised base case analysis and any remaining uncertainty in the context of the rarity of PAH, the considerable unmet need for this group of patients and the potential value of sotatercept to patients and society that it has not been possible to capture within the cost-effectiveness analysis.

1	**Population/positioning in which sotatercept is initiated (DG Section 1.2 and 3.4)** The company have identified two points in the DG where the DG incorrectly states the positioning of sotatercept in this decision problem. 1) Section 1.2 of the DG states: *"The company asked for sotatercept to be considered only for people who have an intermediate–low-risk status to reflect the populations in the clinical trials.*" This is incorrect; the company made a submission for this population based on the available comparative evidence in line with the NICE Manual. The STELLAR trial enrolled patients from all risk status groups. Furthermore, intermediate-low risk is proposed as the appropriate population in which to initiate sotatercept, not as the population to whom use should be restricted. Patients who benefit from treatment such that they are then classified as low risk, or patients who experience progression to intermediate-high risk should also be treated with sotatercept, as the STELLAR trial has demonstrated clinical benefit in these populations. The company requests that this statement be updated as follows: "*The company asked for sotatercept to be considered for initiation in people who have an intermediate–low-risk status based on the available comparative evidence*". 2) Section 3.4 of the DG states: "*The committee acknowledged that this treatment was positioned*

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

	by the company only for PAH with an intermediate–low-risk status and so would not fully address the unmet need." The company requests that this statement be updated to reflect the positioning of the treatment as **initiation** in intermediate-low risk.
2	**Comparative efficacy of sotatercept versus selexipag (DG Section 3.7)** Section 3.7 of the DG states: "*STELLAR used ESC/ERS risk stratification guidelines whereas GRIPHON used WHO FC risk stratification".* WHO FC was also a stratification factor in the STELLAR trial. Access to individual patient-level data for the three endpoints which are used to calculate ESC/ERS risk status enables assignment of risk status to STELLAR participants. The company does not have access to patient- level data for GRIPHON (the study was sponsored by Actelion Pharmaceuticals Ltd.), so is unable to retroactively assign risk status to those participants, and hence a trial-based comparison of sotatercept versus selexipag for the endpoints of risk status improvement/worsening is not feasible.
3	**Post-hoc analysis using STELLAR data (DG section 3.8)** "*The committee…noted the potential benefit of further analysis, such as propensity score matching, to help to account for the differences in subgroups' baseline characteristics and to help with understanding how differences may impact the results of this analysis. Such an additional analysis may help to determine to what extent the within-trial analysis data can be used, if at all."* The STELLAR trial was not designed to make a randomised comparison of sotatercept vs. selexipag for the endpoints of change in ESC/ERS risk stratification. Furthermore, indirect comparison of sotatercept vs. selexipag for these endpoints using data from their respective trials was not possible due to change in ESC/ERS risk stratification data not being reported for patients treated with selexipag in those trial that studied selexipag (i.e. the GRIPHON and TRACE trials). The within-trial analysis presented in section 2.8.1 of the Company evidence submission to address this data gap is post-hoc and not a randomised comparison. The population characteristics between two arms in a subgroup analysis of a study being balanced to the same degree as that of the full trial population post systematic randomisation is unlikely. As described in section 2.8.1 of the Company evidence submission, the company identified a list of characteristics that could be important baseline prognostic factors (PVR, BNP/NT-proBNP, 6MWD, and WHO FC. Substantial imbalances in these characteristics between the arms of the post-hoc populations were not observed. The committee noted some differences between the compared groups in the characteristics of age, time from diagnosis, and baseline risk status, however the impact of these differences on the direction or magnitude of the results is uncertain. These may be relevant potential effect modifiers and the Company understands the potential benefit of adjusting for these, if feasible. However, it is not appropriate to perform such a propensity score matching/adjusted analysis in this particular case, for the reasons outlined below. An unbiased treatment comparison requires that, at baseline, the two treatment groups are comparable and if this is not the case, it can be achieved by statistical adjustments, such as propensity weighting (such as that described in NICE technical support document 17 [TSD 17]) (1). As described in TSD 17, matching works best if there are a large number of individuals to use in the matching cohort. To be statistically viable, there is need for a sufficient sample size available for the analysis, especially in the context of propensity score matching analyses which tend to decrease the effective sample size through

	exclusion of patients without a viable match. It is necessary to maintain a sufficient number of patients in the matched/adjusted analysis so that its results can be interpreted meaningfully. However, the more covariates that need to be matched/weighted on, the harder this becomes to achieve. In the case of the within-trial analysis of STELLAR, there are small numbers of relevant patients available to inform the comparison: █ patients in the sotatercept group and █ patients in the selexipag group. With the number of patients already being low prior to any potential matching/adjustment, modelling propensity scores and performing an analysis based on these, is not viable. In order to meaningfully control for the correlation between the treatment effect and even a small number of potentially relevant covariates), a notably larger sample size would be required to enable a methodologically appropriate and meaningful comparison (i.e. one that could produce results with a greater degree of certainty than the existing unadjusted analysis). As a result, the company has not conducted propensity score analysis based on the within-trial analysis. In lieu of conducting a propensity score analysis, the company submission presented two different approaches for the comparison of sotatercept vs. selexipag: (i) the (naïve) within-trial analysis of STELLAR, and (ii) the Bucher ITC using sotatercept WHO FC data from the STELLAR study and selexipag data for the GRIPHON/TRACE studies pooled. However, in response to the committee request to adjust for population differences at baseline, the Company has also conducted a matching-adjusted indirect comparison (MAIC), described in response to DG section 3.10 below. **Short-term transition probabilities for selexipag in the economic model (DG Section 3.10)** *The committee suggested that a matching-adjusted indirect comparison could be used in this ITC. This could address the differences in baseline characteristics between STELLAR and GRIPHON and TRACE and would make the analysis more robust*. In response to the suggestion above, the company has conducted a matching-adjusted indirect comparison (MAIC) comparing sotatercept to selexipag for the endpoints of WHO FC improvement and worsening (see Appendix A of this response document for full details). The systematic literature review performed to identify relevant studies (as originally described in Appendix B of the Company Evidence Submission) was used to inform this MAIC. The systematic literature review identified the GRIPHON and TRACE studies that compared selexipag + background therapy to placebo + background therapy, that could allow anchored ITC against sotatercept using data from the STELLAR study. It is not possible to pool the data from the GRIPHON and TRACE studies together into a single analysis as only aggregate data are published/available from these studies. Individual patient data from these studies are required to create a pooled selexipag arm for the sotatercept data to be matched/adjusted to in MAIC. The GRIPHON study was chosen to be used in the MAIC due to the considerably greater sample size of the GRIPHON study compared to the TRACE study (N=1,156 versus N=108, respectively) and the greater comparability of the baseline characteristics of the GRIPHON study population to the STELLAR study population, compared to than that of the TRACE study population (as described in section 2.10.1 and Appendix P of the company evidence submission). A key difference between the STELLAR and GRIPHON trials is the distribution of background therapies patients were receiving at randomisation. A proportion of STELLAR patients were receiving background triple therapy, whereas all patients in the GRIPHON trial were receiving background monotherapy or dual therapy – a reflection of the evolution of the standard of care in the time that has elapsed since the publication of the GRIPHON trial. It is stated in section 3.8 of the DG: "*People who are having triple*

	combination therapy are having more intensive treatment and so are likely to have a worse prognosis.", reflecting the committee's concerns that the inclusion of STELLAR background triple therapy patients may impact the comparability of the trials and hence the validity of the results. To address this, a STELLAR subpopulation was chosen for the MAIC analyses: the participants who were on PAH background monotherapy or dual therapy. This subpopulation was then reweighted to further improve the baseline comparability of the STELLAR and GRIPHON populations. The following patient characteristics at baseline were used as pre-specified potential treatment effect modifiers (identified through literature review and clinical insight): • WHO diagnostic group • WHO FC • Age • Six-minute walking distance. These align with the list of potentially important baseline prognostic factors identified previously (see section 2.8.1 of the Company evidence submission). However, a shorter list was used for this analysis. For the MAIC, matching larger numbers of average baseline characteristics and adjusting for greater cross-trial baseline differences would require more extreme weights and reduce the effective sample size (2). It is preferrable to keep the effective sample size of the analysis high in order for it to have high statistical power, so only covariates with the highest likelihood of being the most important treatment effect modifiers were prioritised. Furthermore, baseline data for all factors were not available from GRIPHON. As with the Bucher comparison, ESC/ERS risk status improvement/worsening were not feasible endpoints for analysis due to insufficient data from the selexipag trials, but comparison in terms of WHO FC improvement/maintenance/worsening was feasible. The MAIC results are summarised below (and described in more detail in Appendix A of this response form) in terms of the risk ratio (95% confidence interval) of improvement/maintenance/worsening from baseline in WHO FC at Week 24 (STELLAR)/Week 26 (GRIPHON), of sotatercept + background therapy versus selexipag + background therapy: • WHO FC improvement: [redacted] • WHO FC maintenance: [redacted] • WHO FC worsening: [redacted] The results show that treatment with sotercept is significantly less likely to lead to WHO FC worsening than treatment with selexipag. The results for the relative risk of WHO FC improvement and maintenance are not statistically significant, although the point estimates for each of these favour sotatercept. The precision of the MAIC results (i.e. how narrow the 95% confidence intervals are) is reduced compared to that of the Bucher ITC analysis due to the effective sample size of the MAIC being lower after matching, and the sample sizes from STELLAR and GRIPHON differ notably (as presented in Appendix A of this response). However, the point estimates are presented as a more appropriate measure of the treatment effect of sotatercept vs. selexipag for the endpoints of WHO FC improvement and worsening than the Bucher ITC, addressing the committee's concerns about the comparability of the patient populations. Previously, the ITC results provided were based on a comparison between different populations that were not adjusted to reflect that. As a result, the revised company base case uses the MAIC point

	estimates to derive short-term transition probabilities for selexipag in the economic model.
4	**Modelling approach (DG Section 3.9)** The DG outlines the company and EAG approaches to modelling, focusing on the model structure and approaches to modelling mortality in the economic model. This response provides clarification to address inaccuracies in the DG, and further justification for the company-preferred approach to modelling mortality, which is also preferred in the EAG base case. The response is structured as follows: 1) Clarification on differences between modelling approaches for sotatercept and selexipag 2) Clarification on the company and EAG approach to modelling mortality in the economic model 3) Justification for company and EAG approach to modelling mortality ***Clarification on differences between modelling approaches for sotatercept and selexipag*** The DG states that "*The EAG raised concerns around differences in how sotatercept and selexipag treatment is started (see section 3.12)",* which the company does not understand to be the case. The company request that this statement is updated to *"The EAG raised concerns around differences in how sotatercept and selexipag treatment are discontinued upon progression to the intermediate-high or high risk health states, and the subsequent initiation of PG12 analogues (see section 3.12)."* The rationale for this amendment is outlined below. The company's model does not include any differences between starting treatment with sotatercept and selexipag and we are unaware of concerns raised in the EAG report or subsequently with regards to the company approach for starting either treatment in the economic model. The company approach follows the ESC/ERS guidelines, which specify treatment combinations per the 4-strata risk status. In the company's model, both sotatercept and selexipag are initiated in intermediate-low risk on top of dual background therapy (PDE5i + ERA), an approach which is replicated in the EAG base case. The cross reference in the DG to "(see section 3.12)" refers to subsequent initiation of PG12 analogues upon progression to intermediate-high and high-risk health states, and this was also the subject of discussion in the committee meeting. ***Company and EAG base case approach to modelling mortality*** Section 3.9 of the DG incorrectly states that "*The company used a dependent parametric regression model (Gompertz) with non-low-risk states fitted as covariates fitted to the Kaplan–Meier survival data.*" MSD request that this statement is updated to "*The company fitted one parametric distribution (Gamma) to the low-risk health state Kaplan-Meier survival data as a reference state and applied a hazard ratio for non-low-risk health states.*" Uncertainty in the company approach to modelling mortality is also discussed and the committee requested additional justification and scenario analysis. A summary of the mortality approach applied in both the company and EAG base case is presented below. • One single parametric distribution (Gamma) fitted to low risk, HR applied for non-low-risk health

	states. As described in the company submission Section 3.3.2, the Guyot algorithm was used to reconstruct individual patient-level data from published Kaplan-Meier curves stratified by risk stratus from Rosenkranz et al. (2023) and survival functions were fitted to the low-risk strata. HRs for non-low risk health states were similarly obtained from Rosenkranz et al. (2023). o This aligns with previous mortality modelling in a WHO FC-based model in PAH (3) where a FC I parametric distribution has been fitted with HR applied to derive mortality curves for the FC II-IV health states Both the company and EAG selected the hazard ratio approach with a Gamma distribution for low risk as their base case. An alternative approach to modelling mortality was considered in the company model, and is presented in the scenario analysis in the revised model. In this approach, a parametric regression model was jointly fitted to all available Kaplan-Meier survivor data for each risk strata, reported by Rosenkranz et al (4), using non-low risk health states as covariates. As with the base case approach, the choice of distribution was based on hazard patterns and clinical plausibility. The Gompertz and Weibull models showed clear monotonically increasing hazard patterns which were in line with clinical expert expectations, and the Gompertz approach was selected as the most suitable under the dependent mortality approach. This is the model selected for the scenario analysis which follows the dependent mortality approach. While fitting one overall model for capturing mortality would typically be preferred, in this case, we opted for a simpler base case model using the low-risk group as a reference curve, to ensure that the model inputs can be more readily understood and explored through sensitivity analysis. While both approaches are subject to limitations, the hazard ratio approach does not rely on estimating absolute survival outcomes for each risk group, which introduces challenges for survival extrapolation (e.g., the high risk Kaplan-Meier curve is informed by only 25 patients at baseline, which reduces to fewer than 10 patients by two years). Consequently, the single parametric mortality approach was concluded to better capture the uncertainty in long-term survival in the company base case, and was also selected by the EAG for their base case. Nevertheless, the dependent, jointly-fitted, approach is presented in a scenario analysis to demonstrate the impact on model results, as requested in the DG. It should be noted that, compared to the base case approach, the dependent mortality approach using Gompertz jointly-fitted to all risk states results in a higher hazard rate of mortality over time in each of the risk states. This has a greater impact for patients in the selexipag arm, who progress to the higher-risk health states more quickly in the model due to inferior efficacy, and hence assume higher risk of mortality compared to the sotatercept arm. As such, the company and EAG selected base case approach is more conservative with regards to the comparative survival of sotatercept versus selexipag.
5	**Long-term transition probabilities (DG Section 3.11)** In the DG, the committee expressed their preferred assumption for long-term transition probabilities for selexipag to be "*derived by applying half of the relative risk reduction of disease progression for sotatercept versus selexipag at 24 weeks, based on the ITC with GRIPHON and TRACE*" whilst also requesting additional data to support validation of the relative risk reduction.

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

	As per the discussion in comment 4 above, the committee suggested that a MAIC may be more appropriate than the unadjusted ITC, and the company has subsequently conducted this analysis. The methodology and results have been discussed above. For consistency between the short-term and long-term transition probabilities, the company applies the same relative risks for improvement/worsening between sotatercept versus selexipag in the revised company base case i.e. those estimated by the MAIC. The EAG applied a range of relative risk reductions when estimating the long-term transition probabilities. The company views this range as arbitrary and implicitly reflective of a wane in sotatercept relative efficacy which has not been observed in the open-label long-term follow-up SOTERIA trial. The committee discussed the need for alternative data sets to inform and validate the relative risk reduction. The company has consulted the literature for data reporting risk status distribution over time of intermediate-low risk patients initiated with selexipag, who would be the most generalisable population to the modelled cohort. An appropriate and available data set was not identified. As discussed in the company evidence submission, the company attempted to analyse real-world evidence reporting the risk status distribution of patients over time following their initiation of selexipag from a UK-based registry, however sufficient sample size was not available for analysis (see Appendix O of company evidence submission). In the absence of suitable data sets to establish the real-world projections for selexipag, the company has revised the base case to model long-term transition probabilities for selexipag as follows: • Application of relative risks for risk status improvement/worsening as estimated by the MAIC vs. GRIPHON (uses WHO FC improvement/worsening as a proxy) • No reduction in relative risk of risk status worsening applied to model cycles beyond week 24. The company acknowledges the clinical expert comments in the committee meeting and given the absence of suitable external data and difficulties in comparing to the primary endpoint of GRIPHON, the uncertainty around this model assumption is explored further in scenario analysis. The company acknowledges the subject of relative efficacy between sotatercept and selexipag in the long-term as an area of much uncertainty, due to the absence of head-to-head data and differences in trial design and comparability of populations. Given the rarity of PAH and the significant need for innovation in this disease area, the company request the committee to accept a higher degree of uncertainty in decision making for this treatment.

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

6	**Initiation of PGI2 analogues in the sotatercept arm (DG section 3.13)** The company have revised its base case to include the initiation of PCA in 39.9% of patients who progress to intermediate-high risk in the sotatercept arm, in line with the committee-preferred assumption. As per the committee request to include the potential for patients receiving PCA to improve risk status discussed above, the capacity to improve has been assigned to this proportion of intermediate-high risk patients receiving the combination of sotatercept + PCA, but not to the proportion receiving sotatercept without PCA, through the model averaging approach described in company response 7 below. The company and EAG base cases both included the assumption of stopping sotatercept upon progression to high risk due to lack of data in STELLAR for transitions from high risk when sotatercept is added to PCA. In response to the committee request to explore other data sources, the company explored data from the ZENITH trial (5). In contrast to STELLAR, the sotatercept arm in the ZENITH trial included some high risk patients (██ at baseline), and the company investigated the feasibility of informing a comparison between sotatercept added to background PDE5i+ERA+PCA compared to PDE5i+ERA+PCA (the standard of care) in high risk patients. Examination of the ZENITH trial data was conducted to understand how many of these high risk patients were receiving the appropriate background triple therapy regimen to inform this comparison. A limited number (██) were receiving PDE5i+ERA+any form of PCA. It was also assessed if a transition probability was estimable for these patients i.e. a recorded risk status measurement at baseline and at week 24, and the sample size was even lower. In the ZENITH SoC arm, relevant patient numbers were even more limited, with █ high risk patients at baseline, █ of whom were receiving PDE5i+ERA+any form of PCA. The SoC arm in particular experience attrition over the trial follow-up period, with just █ high risk patients having an estimable transition probability. It is unknown if any of these patients were receiving the appropriate background regimen for this comparison. Based on this assessment, the sample size was determined to be too low and uncertain to present transition probabilities for the addition of sotatercept to PDE5i+ERA+PCA (the standard of care) in high risk patients, and also retained the assumption in the model base case that patients discontinue sotatercept upon entry to the high risk state (with the exception of the 15% of patients who are not eligible or accepting of a PCA, as per the committee-preferred assumption).
7	**No clinical improvement after PGI2 initiation (DG Section 3.12)** In the DG, the committee requested further analysis requiring an adaptation to the current economic model structure to allow for clinical improvement after starting PCA: "*the model structure should reflect improvements in risk status after starting PCA*". However, it is acknowledged by clinical experts in the DG that this capacity to improve is only applicable to a subset of those who initiate the treatments: "*The clinical expert agreed that clinical improvement can be seen after starting intravenous PCAs. But, they stated that the main benefit of intravenous PCAs is stability in the intermediate–high or high-risk states.*" In response to the committee request, the company has updated the model structure. The following health states have been added for both model arms: • An intermediate-high risk tunnel state where patients spend one cycle following their

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

	progression from intermediate-low risk: this state is necessary to distribute the patients to subsequent health states, reflecting the statement above that PCA initiation has a range of clinical outcomes following initiation. Furthermore, it avoids the possibility of infinite looping of patients through the health states, if patients were routed to the current health states as per the programming of the originally submitted model. The company consulted clinical experts (N=4) about how soon a clinical improvement could be detected. The feedback varied depending on patient type, but it was agreed that for the average patient, it is reasonable to assume a response to PCA would occur within 12 weeks of initiation. • An intermediate-low risk state (history of or on PCA), where patients who initiated PCA can improve to from the tunnel state: clinical expert advice confirmed that these patients would remain on PCA once they improve. This state is necessary to distinguish these intermediate-low risk patients from those who have never worsened and continue to be treated with selexipag. • An intermediate-high risk state (history of or on PCA), where patients who maintain their status following initiation of PCA move to from the tunnel state. • A high risk state (history of or on PCA), where patients whose status worsens following initiation of PCA move to from the tunnel state. A number of assumptions were made to enable the updated structure to retain a logical and manageable patient flow: • Patients initiating PCA are offered a one-cycle opportunity to improve back to intermediate-low risk • Further improvement beyond intermediate-low risk to low risk is not modelled • Patients who enter the intermediate-high risk state (history of or on PCA) do not have the capacity to improve, they can remain in intermediate-high or worsen to high risk • Patients who enter high risk do not have the capacity to improve. As stated previously by the company, there is a paucity of suitable data to inform transitions between these health states. However, the company has made best use of the available evidence to inform a cohort Markov model, in order to comply with the committee's request. The transition probabilities used in the updated structure are outlined below. *Selexipag arm* Transitions from the intermediate-high risk tunnel state (where patients are treated with a mixture of PCA [85% of patients] and selexipag [15% of patients], as per the committee-preferred assumption discussed later) are informed by the SoC arm from the STELLAR trial. Week 12-24 values were selected for consistency with the wider model. This arm, whilst not aligning with the model health state on exact proportions of patients receiving the treatments, represents the efficacy of a patient group receiving a combination of non-sotatercept treatments, including selexipag, inhaled and IV PCA.

Table 1: Transition probabilities for PCA improvement in the selexipag arm

To: From:	Intermediate-low risk (history of/on PCA)	Intermediate-high risk (history of/on PCA)	High risk (history of/on PCA)	Source
Intermediate-high (tunnel state)				STELLAR SoC arm week 12-24
Intermediate-low risk (history of/on PCA)				STELLAR SoC arm week 12-24
Intermediate-high risk (history of/on PCA)				STELLAR SoC arm week 12-24
High risk (history of/on PCA)				Assumption

Abbreviations: PCA, prostacyclin analogue; SoC, standard of care
*sum of IL > IH and IL > H i.e. 0.15 + 0.017
**sum of IH > IL and IH > IH i.e. 0.12 + 0.72

Sotatercept arm

Upon progression to intermediate-high risk, a proportion of patients receive sotatercept + PCA (39.9%), as per the committee-preferred assumption discussed in the row above. The remaining 60.1% continue to receive sotatercept in combination with PDE5i + ERA. The potential to improve risk status is therefore available to this former proportion of patients. The restriction of no improvement continues to apply to the patients who do not initiate PCA, consistent with the original base case. In order to reflect the movement of these two patient types in the model, two scenarios are modelled:

- 100% of the patients moving to intermediate-high risk receive PCA in addition to sotatercept (all patients are allocated this cost and the potential to improve risk status)
- 0% of the patients moving to intermediate-high risk receive PCA in addition to sotatercept (no patients are allocated this cost and there is no potential to improve risk status)

The total costs and QALYs for each of these scenarios are estimated and then weighted in a ratio of 39.9%:60.1% and the resulting values for this cohort of sotatercept patients who progress to intermediate-high risk are used for the ICER calculation. This model averaging approach was deemed to be more appropriate than attempts to weight transition probabilities live within the patient flow calculations, which may have led to invalid or non-intuitive values owing to structural limitations of a cohort-level model.

Transitions from the intermediate-high risk tunnel state for patients receiving sotatercept + PCA are informed by the sotatercept arm from the STELLAR trial (week 12-24 values were selected for consistency with the wider model). Consistent with the original company base case, transitions from

FCIII were used as proxies for transitions from intermediate-high risk, to overcome the issue of there being zero transitions from intermediate-high to high risk in the STELLAR sotatercept arm as was discussed in the company submission. Transition from intermediate-high risk (history of or on PCA) to high risk hence also uses the FC III to FC IV probability as a proxy.

Risk status transition data, where available, continues to be the preferred source - consistent with the original company base case, transitions from intermediate-low risk (history of or on PCA) to intermediate-high are informed by the STELLAR sotatercept arm (the probability of staying in intermediate-low are the sum of those of staying and improving to low risk, which is not permitted for PCA patients in this structure).

Table 2: Transition probabilities for PCA improvement in the sotatercept arm (Scenario 1: 100% of patients initiate PCA in intermediate-high risk)

To: / From:	Intermediate-low risk risk (history of/on PCA)	Intermediate-high risk (history of/on PCA)	High risk (history of/on PCA)	Source
Intermediate-high risk (tunnel state)				STELLAR sotatercept arm week 12-24
Intermediate-low risk (history of/on PCA)				STELLAR sotatercept arm week 12-24
Intermediate-high risk (history of/on PCA)				STELLAR sotatercept arm week 12-24
High risk (history of/on PCA)				Assumption

Abbreviations: PCA, prostacyclin analogue
*Transitions from WHO FC III in STELLAR sotatercept arm used as a proxy for transitions from intermediate-high risk
**sum of IL >L and IL > IL i.e. 0.292 + 0.625
***sum of IL > IH and IL > H i.e. 0.083 + 0
^sum of FCIII > II and FCIII > III i.e. 0.333 + 0.646

Table 3: Transition probabilities for PCA improvement in the sotatercept arm (Scenario 2: 0% of patients initiate PCA in intermediate-high risk)

To: / From:	Intermediate-low risk (history of/on PCA)	Intermediate-high risk (history of/on PCA)	High risk (history of/on PCA)	Source
Intermediate-high (tunnel state)				STELLAR sotatercept arm week 12-24
Intermediate-low risk (history of/on PCA)				NA

Intermediate-high risk				STELLAR sotatercept arm week 12-24
High risk (history of/on PCA)				Assumption

Abbreviations: NA, not applicable; PCA, prostacyclin analogue
*Transitions from WHO FC III in STELLAR sotatercept arm used as a proxy for transitions from intermediate-high risk

Model schematics for the additional health states in the sotatercept (two scenarios) and selexipag arms are presented in Figure 1, Figure ***3*** and Figure 3 below respectively.

Figure 1: Model schematic for patients who progress from intermediate-low risk on sotatercept (scenario 1)

Figure 2. Model schematic for patients who progress from intermediate-low risk on sotatercept (scenario 2)

Figure 3: Model schematic for patients who progress from intermediate-low risk on selexipag

8	**Proportion of PGI2 analogues (DG section 3.14)** When making the evidence submission, the company sought to reflect the World Symposia guidelines on pulmonary hypertension, which specifies the standard of care in intermediate-high and high risk to be PDE5i+ERA+PCA. The treatment distribution in the company base case reflected this i.e. all patients in these states are treated with the standard of care. The EAG's approach, in which less than 100% of people who progressed received PCA, was supported by the estimates provided by clinical and patient experts during the meeting, and those subsequently consulted by the company. This estimate is used in the revised base case. 85% of patients who progress to intermediate-high and high risk states in the selexipag arm, and to high risk in the sotatercept arm receive PCA with the remainder continuing to receive the parent treatment (either selexipag or sotatercept, respectively). 39.9% of the patients progressing to intermediate-high risk in the sotatercept arm receive PCA in addition to sotatercept, in line with EAG base case and committee preferred assumptions.
9	**Weight-based dosing of intravenous PCA (DG Section 3.15)** In the absence of a previous NICE TA in PAH, it was not possible to identify a previous committee-preferred value for this input. The company conducted a review of HTA reports in other jurisdictions and identified a value reported in the most recent CDA appraisal in PAH (6), which was presented in the company base case. The company did also not identify UK guidelines for the dosing of these medicines A clinical expert consulted by the EAG proposed the ESC/ERS guidelines as a more appropriate reflection of UK clinical practice, and the EAG preferred to use the midpoint values of the ranges

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

	proposed for epoprostenol (23ng/kg/minute) and treprostinil (42.5ng/kg/minute) in their base case. Since the committee meeting, the company consulted a panel of four UK clinical experts, from a range of PAH centres to understand their approach to dosing of the intravenous medicines. The experts agreed that, whilst very high doses are used in certain patients, the company base case value sourced from the CDA report is higher than that used in the average UK patient. However, it was discussed that patients who have previously been receiving a treatment which targets the prostacyclin receptor (e.g. selexipag) are suitable for higher target doses than those who have not, due to reduced likelihood of tolerance issues, and that a range of 30-40ng/kg/minute of epoprostenol is more appropriate for these patients. It was agreed that the EAG-preferred values are appropriate for a selexipag-naïve patient, who has not previously used a treatment which targets the prostacyclin receptor. It was also stated that the dose of treprostinil is typically twice that of epoprostenol, reflecting clinical opinion received by the EAG. Taking this feedback into account, the company applied differential doses of intravenous PCA to treatment arms in the revised base case. Intermediate-high risk patients using epoprostenol and treprostinil following selexipag use in intermediate-low risk were assigned doses of 35ng/kg/minute (midpoint of the range provided) and 70ng/kg/minute (twice that of epoprostenol), respectively. Intermediate-high risk patients using the medicines following sotatercept use in intermediate-low risk were assigned epoprostenol and treprostinil doses of 23ng/kg/minute and 42.5ng/kg/minute, respectively, as per the EAG base case. Upon progression to high risk, sotatercept patients are also escalated to the higher doses to reflect disease worsening and the need for more intensive treatment. Furthermore, "*The committee acknowledged that the company has accounted for wastage in its modelling.*"
10	**Administration disutility in the economic model (DG Section 3.16)** *Hospitalisation-related disutility* The DG states the following committee conclusion: *"the utility decrement associated with hospitalisation should be reduced from 0.105 to 0.071, applied for the duration of the cycle in which events occur (to align more closely with utility for hospitalisation between 0 and 90 days) and if no possibility of clinical improvement after starting PGI2 is added".* The company request that the underlined text be removed as the discussion about hospitalisation utility decrement is unrelated to the issue of PGI2 clinical improvement – this has been conflated with the bullet point below it in the DG. The committee-preferred value for hospitalisation-associated disutility has been included in the revised base case. *Disutility for intravenous PCA administration* In the DG, the committee notes the uncertainty associated with the company economic model by not capturing potential risk status improvement following the initiation of PG12 analogues and as a

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

	consequence the committee-preferred assumption is to not include a disutility for intravenous PCA products. As described in comment 6 above, the company have updated the economic model to allow for improvement in risk status in those patients initiating PCA. In the revised company base case, the disutility associated with IV administration of PCA continues to be applied to reflect the impact of receiving these treatments on patient quality of life, beyond that of simply having PAH. This impact has been described in expert statements received both prior to and during the committee meeting, from clinical and patient experts, as well as in research published by PHA UK (7-9). The company view is that to exclude a disutility for these treatments is to overstate the quality of life of PAH patients receiving IV PCA. It is not feasible to derive a specific disutility value for patients receiving IV PCA from the STELLAR trial for the following reasons: • EQ-5D-5L was an exploratory endpoint in the study, with the goal of evaluating differences in health-related quality of life between the treatment arms in general. As such, the STELLAR trial was not powered to evaluate differences in health-related quality of life between those on- or off-PCA • Only about a quarter of patients in each of the risk states in the study were receiving an IV treatment at baseline so this impact has likely not been accounted for in the health state utility values reported i.e. they are more reflective of a patient not receiving IV than one who is • Given the IV PCA medicines are initiated in the model following other treatments, the relevant disutility value is for a patient moving from off-IV to on-IV. These treatments were administered as background therapies in STELLAR, i.e. the majority of patients did not initiate IV during the study so this within-patient comparison cannot be robustly made. Furthermore, the imbalance in sample sizes between on-IV and off-IV is likely to bias any value that would be estimated. The Davies study reports a disutility difference for a patient receiving an IV treatment relative to an oral treatment such as selexipag (10), which is the transition that patients in the model make upon progression to intermediate-high risk. It is assumed that this disutility also applies to a patient who has also been receiving treatment only with oral tablets and a subcutaneously-administered treatment like sotatercept, where the patient has been taught to self-administer across multiple appointments (as per the EAG base case). This peer-reviewed study elicited disutility values from a robust sample of the UK general public (n=150) with a broad geographical spread, who completed the EQ-5D-5L questionnaire and valued health states differentiated by route of administration of PAH medicines in time trade-off interviews. The health state descriptions they valued were drafted based on a literature review and interviews with three clinical experts, including two PH clinical nurse specialists who, based on the company's knowledge of the UK treatment setting, are highly familiar with the patient experience while receiving PAH medicines. Each health state contained an identical description of the disease, symptoms, and impacts of PAH but differed in the description of the treatment mode of administration. Furthermore, a support document was developed that included pictures and further descriptions of the treatment modes of administration. The health state descriptions presented in this study are robust and comprehensive. The results of the study demonstrated a significant ($p<0.001$) difference between the oral and intravenous health state descriptions when valued by the general population sample. A disutility was also reported for the inhaled route of administration, however in the company base case, it was

conservatively assumed that the patients receiving inhaled iloprost do not experience a disutility, which offsets some of the inevitable uncertainty associated with these HRQoL inputs.

The revised company base case continues to apply the disutility reported by Davies to patients receiving IV PCA, whilst also including the capacity to improve risk status with these medicines as requested by the committee. To explore the uncertainty around this input, two additional scenarios were investigated.

- Despite the fact that the health states valued by the Davies respondents included identical descriptions of the disease factors (i.e., the disutility refers exclusively to the impact of administration route), the company considered the possibility of double counting of disutility between worsening from intermediate-low to intermediate-high risk and initiation of IV PCA. Patients in the model who make this transition experience a disutility of ████ (difference in HSUV between intermediate-low vs. intermediate-high risk). A proportion of these patients also initiate IV PCA. The company base case applies both reductions in HRQoL to these patients, reflecting the independent effects on HRQoL of these events. To investigate the impact of potential double counting of disutility with this approach, a scenario analysis applies the net disutility of ████ (0.307 minus ████) to patients initiating and receiving IV PCA.
- The EAG raised concerns that the company approach underestimates the HRQoL of patients receiving IV PCA, as the potential improvement of PAH risk status (and hence HRQoL) with these medicines is excluded, i.e. that the reductions in HRQoL due to the administration route could be offset by improvements due to improved risk status. With the revised model structure described above, this assumption is explored further. The company propose that the offset can only apply to patients who do indeed receive a risk status improvement. In the intermediate-high risk tunnel state (first model cycle on IV PCA), all IV PCA patients receive the Davies disutility. Any patient who does not experience the improvement back to intermediate-low risk stays in intermediate-high risk and continues to receive the Davies disutility due to no improvement-related offset. The patients who do improve on PCA receive the reduced disutility value of ████ as long as they remain in the intermediate-low risk state, to reflect that the impact of receiving an IV treatment has a lesser impact on their HRQoL because it has simultaneously improved their PAH i.e. they receive a “double increment”: an improvement in HSUV due to moving to a better health state and also a reduction in the administration route-related disutility.

Table 4: Summary of base case and scenario analysis inputs for IV disutility

Health State	IV PCA disutility			
	Company base case	**Scenario 1**	**Scenario 2**	**EAG base case**
Low Risk	N/A	N/A	N/A	N/A
Intermediate-Low Risk off PCA	N/A	N/A	N/A	N/A
Intermediate-Low Risk (History of/on PCA)	-0.307	████	████	N/A
Intermediate-High Risk PCA tunnel state	-0.307	████	-0.307	0

	Intermediate-High Risk | -0.307 | ███ | -0.307 | 0 High Risk | -0.307 | ███ | -0.307 | 0 *Abbreviations: EAG, evidence assessment group; IV, intravenous; N/A, not applicable; PCA, prostacyclin analogue*
11	**Severity Modifier (DG Section 3.18)** PAH is a severe and debilitating condition, which causes people to forego substantial health over a lifetime. This has been underscored by the expert statements received both prior to and during the committee meeting, from clinical and patient experts. In both the company and EAG base case, the threshold for qualification for a 1.2 severity modifier for QALY gains is met, based on absolute shortfall. As noted in the DG, there are a limited number of scenarios where the severity modifier does not apply, including where the model uses baseline characteristics reported for adult patients in the longitudinal cohort (2009-19) in the 2019 UK National Audit of Pulmonary Hypertension (NAPH) – the most recent audit version to report demographic data for PAH patients. The mean age from this data source (58.0 years) is older than the mean age from the STELLAR trial (47.9 years), which was used in the company and EAG base cases. The data reported for the NAPH cohort refer to the entire population in Great Britain treated for PAH (2009-19), which is not a generalisable cohort to the population under consideration in this appraisal i.e. intermediate-low risk patients receiving background dual therapy. The NAPH cohort is notably broader, and the STELLAR trial mean age is more appropriate for determining a weighting to apply to QALYs generated using data from the STELLAR population. It is not possible to reliably adjust the STELLAR transition probabilities or utility scores to a hypothetically older population, who are likely to be receiving a range of different standard of care treatments. Referring to the subpopulation of patients in the trial who received sotatercept on top of PDE5i + ERA (██), a mean age of 46.0 years was observed, indicating that the subpopulation closest to the anticipated population in clinical practice may be younger than the average patient in the trial. Given the patient population under consideration, the company maintain their position that the appropriate characteristics on which to assess the severity of this disease are those used in the company and EAG base cases i.e. the mean age from the STELLAR trial. A scenario analysis is presented where the NAPH baseline characteristics are used is presented as requested. It should be noted that in the updated company base case model, the scenario using the NAPH baseline characteristics still meets the threshold qualifying for a 1.2 severity modifier. Different methods of modelling mortality have been explored and are described in more detail in comment 5 above. The company reiterate their original base case approach of using hazard ratios relative to low risk derived from a study reporting survival per risk stratum (4), rather than a dependent model approach, as described above. The company note that the mean age of the population in the study which informed mortality, an input which is linked to age, is 51.7 years i.e. more closely aligned to the STELLAR mean, so using the older NAPH value adds further uncertainty to the survival of the model cohort. Given the profile of PAH and the population it affects, it would be counterintuitive for a severity modifier to not apply in this appraisal.
12	**Benefits not captured in the QALY (DG Section 3.21)** Section 3.21 of the DG states the committee: "*did not identify additional benefits of sotatercept not*

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

	captured in the economic modelling. So the committee concluded that all additional benefits of sotatercept had already been taken into account." However, a considerable list of additional benefits of sotatercept not captured in the QALY calculation are described in detail in section 3.13 of the Company evidence submission, including benefits in terms of: • Patient and caregiver productivity • Increased capacity on transplant list • Psychological benefit of an innovative treatment option • The potential for stable disease to lead to more remote management of PAH • Wider cardiovascular health benefits from improved exercise capacity • Reduced use of anticoagulation medicines • Insensitivity of utility measure to disease changes in PAH • Survival benefit beyond time spent in improved health states • Selexipag adverse event impact on quality of life The company request that these additional benefits be considered by the committee in their selection of preferred base case settings and willingness-to-pay threshold.
13	**Conclusion (DG section 3.22)** The committee concluded that *"Further analysis is needed for:* *• comparing sotatercept with PCA analogues in the intermediate–high and high-risk populations, using published data discussed by clinical experts in the meeting".* In the company base case, sotatercept is administered to intermediate-high risk patients, following progression there from intermediate-low risk, where the treatment was initiated. The company has concluded that there is still no feasible comparison with PCA analogues – the conclusions of the ITC feasibility assessment remain as in the company evidence submission. A feasibility assessment for indirect comparison of sotatercept versus PCA analogues, informed by a systematic literature review to identify potentially relevant evidence, was conducted and appropriate indirect treatment comparisons were not found to be feasible for this comparison using available data (as described in section 2.10.3 and Appendix Q of the Company evidence submission). Section 3.22 of the DG also states the committee "*noted that while the company positioned it for people with intermediate–low-risk status, there is evidence published since the company submission that suggests this could benefit people in intermediate–high or high-risk states."* As stated above, the company proposed a position where sotatercept is initiated in intermediate-low risk and continued in intermediate-high risk. Since the company submission, the ZENITH trial has been published (5), which continues to demonstrate the clinical benefits of sotatercept, in a population consisting of WHO FC III and IV patients. This trial includes patients with intermediate–high or high-risk PAH, confirming the benefit which the DG refers to, however this trial, like STELLAR, is a placebo-controlled study, so comparisons versus PCA analogues are not informed by the trial, and the infeasibility of an ITC in intermediate-high risk patients remains.

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

	In contrast to STELLAR, the sotatercept arm in the ZENITH trial included some high risk patients (█ at baseline), and the company investigated the feasibility of informing a comparison between sotatercept added to background PDE5i+ERA+PCA compared to PDE5i+ERA+PCA (the standard of care) in high risk patients – a comparison which was not possible using STELLAR data. Examination of the ZENITH trial data was conducted to understand how many of these high risk patients were receiving the appropriate background triple therapy regimen to inform this comparison. █████ were receiving PDE5i+ERA+any form of PCA. It was also assessed if a transition probability was estimable for these patients i.e. a recorded risk status measurement at baseline and at week 24, and the sample size was even lower. In the ZENITH SoC arm, relevant patient numbers were also limited, with █ high risk patients at baseline, █ of whom were receiving PDE5i+ERA+any form of PCA. The SoC arm in particular experienced attrition over the trial follow-up period, with just █ high risk patients having an estimable transition probability. It is unknown if any of these patients were receiving the appropriate background regimen for this comparison. Based on this assessment, the sample size is small to inform transition probabilities that would be considered appropriate for decision making. In the absence of appropriate data, the company also retained the assumption in the revised base case that patients discontinue sotatercept upon entry to the high risk state (with the exception of the 15% of patients who are not eligible for or accepting of a PCA, as per the committee-preferred assumption).
14	Please note the company has updated their base case results, and these are presented in Appendix B. The base case has been revised as follows: • Short-term transition probabilities for selexipag: use MAIC vs. GRIPHON point estimates for relative risk of risk status improvement/worsening • Long-term transition probabilities for selexipag: application of MAIC vs. GRIPHON point estimates for relative risk of risk status improvement/worsening, with no reduction in treatment applied • Model structure updated to allow risk status improvement for patients who initiate PCA • PCA initiation in the sotatercept arm: 39.9% of patients receive PCA when they progress to intermediate-high risk • Proportion of PCA use: 85% of patients in intermediate-high risk and high risk in the selexipag arm, and in high risk in the sotatercept arm receive PCA. The remaining patients remain on their initial treatment. • IV PCA dose: EAG doses for epoprostenol and treprostinil applied to patients initiating in sotatercept intermediate-high risk state, higher doses of 35ng/kg/min and 70ng/kg/min, respectively, used in selexipag intermediate-high risk state. Equivalent dosing in high risk state across arms. • Hospitalisation-related disutility: updated to EAG value • Hospitalisation HR between sotatercept and selexipag set equal to 1 • Exclusion of drug wastage costs for sotatercept

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

	Each of these changes has been made in an editable copy of the economic model, so that all changes can be verified against the originally submitted results, and both the company's and the EAG's original base-case analyses can be re-estimated within the same file, for transparency. Appropriate sensitivity and scenario analyses are also provided in the appendix. Please note guidance for the EAG on the implementation of the new model structure and inputs are provided in Appendix C of this document.

Insert extra rows as needed

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about funding from the company and links with, or funding from, the tobacco industry.
- Combine all comments from your organisation into one response. We cannot accept more than one set of comments from each organisation.
- Do not paste other tables into this table – type directly into the table.
- In line with the NICE Health Technology Evaluation Manual (sections 5.4.4 to 5.4.21), if a comment contains confidential information, it is the responsibility of the responder to provide two versions, one complete and one with the confidential information removed (to be published on NICE's website), together with a checklist of the confidential information. Please underline all confidential information, and separately highlight information that is submitted as 'confidential ███' in turquoise, and all information submitted as '███████████' in pink. If confidential information is submitted, please submit a second version of your comments form with that information replaced with asterixis and highlighted in black.
- Do not include medical information about yourself or another person from which you or the person could be identified.
- Do not use abbreviations.
- Do not include attachments such as research articles, letters or leaflets. For copyright reasons, we will have to return comments forms that have attachments without reading them. You can resubmit your comments form without attachments, it must send it by the deadline.
- If you have received agreement from NICE to submit additional evidence with your comments on the draft guidance document, please submit these separately.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

References

1. Faria R, Alava M, Manca A, Wailoo A. NICE DSU TECHNICAL SUPPORT DOCUMENT 17: THE USE OF OBSERVATIONAL DATA TO INFORM ESTIMATES OF TREATMENT EFFECTIVENESS IN TECHNOLOGY APPRAISAL: METHODS FOR COMPARATIVE INDIVIDUAL PATIENT DATA. 2015.
2. Signorovitch JE, Sikirica V, Erder MH, Xie J, Lu M, Hodgkins PS, et al. Matching-Adjusted Indirect Comparisons: A New Tool for Timely Comparative Effectiveness Research. Value in Health. 2012;15(6):940-7.
3. Thenappan T, Shah SJ, Rich S, Tian L, Archer SL, Gomberg-Maitland M. Survival in pulmonary arterial hypertension: a reappraisal of the NIH risk stratification equation. Eur Respir J. 2010;35(5):1079-87.
4. Rosenkranz S, Pausch C, Coghlan JG, Huscher D, Pittrow D, Grunig E, et al. Risk stratification and response to therapy in patients with pulmonary arterial hypertension and comorbidities: A COMPERA analysis. J Heart Lung Transplant. 2023;42(1):102-14.
5. Humbert M, McLaughlin VV, Badesch DB, Ghofrani HA, Gibbs JSR, Gomberg-Maitland M, et al. Sotatercept in Patients with Pulmonary Arterial Hypertension at High Risk for Death. N Engl J Med. 2025;392(20):1987-2000.
6. CADTH. Selexipag (Uptravi®). Canadian Agency for Drugs and Technologies in Health; 2016 October 2016.
7. PHA UK. What it means to live with PH today2017. Available from: https://www.phauk.org/research-survey-work/living-ph-report/.
8. PHA UK. The true emotional impact of pulmonary hypertension2019 2024. Available from: https://www.phauk.org/living-with-pulmonary-hypertension/the-emotional-impact-of-ph-survey-results/.
9. PHA UK. Breathless, not voiceless: What it means to live with PH today2023 2024. Available from: https://www.phauk.org/research/pha-uk-led-research/breathless-not-voiceless-what-it-means-to-live-with-ph-today/.
10. Davies EW, Llewellyn S, Beaudet A, Kosmas CE, Gin-Sing W, Doll HA. Elicitation of health state utilities associated with the mode of administration of drugs acting on the prostacyclin pathway in pulmonary arterial hypertension. Patient preference and adherence. 2018:1079-88.

Appendix A Matching adjusted indirect comparison details

A1 Objectives

This appendix provides more details of the matching adjusted indirect comparison (MAIC) statistical analyses performed by MSD on the comparison with sotatercept versus selexipag in adult participants with pulmonary arterial hypertension (PAH), using data from the STELLAR and GRIPHON studies.

These MAIC analyses are performed because no head-to-head randomised controlled trials design specifically to compare between sotatercept and selexipag are available.

A1.1 Studies considered for use in this analysis

A summary overview of studies considered for inclusion in this analysis is presented in Table 5:.

Table 5: Summary of studies considered for inclusion in this analysis

Trial	Database Cutoff date	Type of data	Intervention of Interest	Population
STELLAR	December 06, 2022	IPD	sotatercept	Eligible patients had a confirmed diagnosis of pulmonary arterial hypertension (idiopathic, heritable, drug-induced, connective-tissue disease–associated, or after shunt correction) in WHO functional class II or III, excluding subtypes associated with portopulmonary disease, schistosomiasis, human immunodeficiency virus infection, or veno-occlusive disease.
GRIPHON	October 01, 2014	AgD	Selexipag	Participants 18 to 75 years of age who had idiopathic or heritable pulmonary arterial hypertension or pulmonary arterial hypertension associated with human immunodeficiency virus infection, drug use or toxin exposure, connective tissue disease, or repaired congenital systemic-to-pulmonary shunts.
TRACE	February 20, 2020	AgD	Selexipag	Patients aged 18 to 75 years with PAH (idiopathic, heritable, drug and toxin induced, or associated with connective tissue disease, HIV infection, or corrected congenital heart disease [simple systemic-to-pulmonary shunts ≥ 1 year after repair]) diagnosed by right heart catheterization were eligible. Patients were required to have a 6-min walk distance (6MWD) ≥ 100 m and be in WHO FC II to III, without hospitalization or worsening WHO FC in the 30 days prior to screening. Patients had to be receiving an endothelin receptor antagonist alone or in combination with a phosphodiesterase type 5 inhibitor or soluble guanylate cyclase stimulator for ≥ 90 days and at a stable dose for ≥ 30 days prior to randomization. Participation in an exercise-based rehabilitation program was not permitted in the 8 weeks prior to study start or at any point during the study.

Table 6 shows an overview of key study characteristics for STELLAR and potential MAIC comparator studies (GRIPHON and TRACE).

Table 6: Overview of Key Study Characteristics for STELLAR, GRIPHON and TRACE

Study Characteristics	STELLAR (NCT04576988)	GRIPHON (NCT01106014)	TRACE (NCT03078907)
Study design	International multicenter, double-blind, randomized, placebo-controlled, phase 3 study	International multicenter, double-blind, randomized, placebo-controlled, phase 3 study	International, multicenter, double-blind, Placebo-controlled Phase 4 Study
Treatment arm	• sotatercept+SoC • Placebo+SoC	• Selexipag±SoC • Placebo±SoC	• Selexipag+SoC • Placebo+SoC
Year of enrollment	2021 to 2022	2009 to 2013	2017 to 2019
Key inclusion criteria	• PH Group 1 • Age ≥18 years • 6MWD ≥ 150 m and ≤ 500 m repeated twice at screening • WHO FC II or III • On stable doses of background PAH therapy and diuretics for at least 90 days prior to screening; for infusion prostacyclins, dose adjustment within 10% of optimal dose was allowed per medical practice	• PH Group 1 • PAH belonging to the following subgroups of Group 1 (idiopathic, or heritable, or drug or toxin induced, or associated [APAH] with connective tissue disease, congenital heart disease with simple systemic-to-pulmonary shunt at least 1 year after surgical repair, or HIV infection) • Age 18-75 years • 6MWD ≥ 50 m and ≤ 450 m repeated twice at screening	• Symptomatic pulmonary arterial hypertension (PAH) belonging to one of the following subgroups only: ○ Idiopathic ○ Heritable ○ Drug or toxin induced ○ Associated with one of the following: connective tissue disease; HIV infection; corrected simple congenital heart disease. • Male or female between 18 and 75 years old inclusive. • Women of childbearing potential must have a negative serum pregnancy test at planned visits and use an acceptable method of birth control from screening up to 30 days after study treatment discontinuation. • With the following hemodynamic characteristics assessed by right heart catheterization (RHC) prior to randomization: ○ Mean pulmonary artery pressure (mPAP) ≥ 25 mmHg ○ Pulmonary vascular resistance (PVR) ≥ 240 dyn•sec/cm5 (or 3 Wood Units) ○ Pulmonary artery wedge pressure (PCWP) or left ventricular end-diastolic pressure (LVEDP) ≤ 15 mmHg. • Treatment with an endothelin receptor antagonist (ERA) for at least 90 days and on a stable dose for 30 days prior to randomization. • If an ERA is given in combination with a phosphodiesterase-5 (PDE-5) inhibitor or soluble guanylate cyclase (sGC) stimulator, these treatments must be ongoing for at least 90 days and on a stable dose for 30 days prior to randomization. • WHO functional class (FC) II or III at randomization • 6-minute walk distance (6MWD) ≥ 100 m at screening. • Ability to walk without a walking aid. • Valid baseline data for daily life physical activity (DLPA) and PAH-SYMPACT®.
Key exclusion criteria	• PH Groups 2 to 5 • Diagnosis of the following PAH Group 1 subtypes: HIV-associated PAH and PAH associated with portal hypertension. Exclusions in PAH Group I should also include schistosomiasis APAH and pulmonary veno-occlusive disease	• PH Groups 2 to 5 • Patients who have received prostacyclin or its analogs within 1 month before baseline visit, or are scheduled to receive any of these compounds during the trial	• Patients on a PAH-specific monotherapy targeting the nitric oxide pathway (i.e., PDE-5 inhibitor or sGC stimulator). • Patients treated with prostacyclin, prostacyclin analog or selexipag within 3 months prior to screening. • Any hospitalization during the last 30 days prior to screening. • Severe coronary heart disease or unstable angina. • Documented severe hepatic impairment or severe renal insufficiency at screening.

			• Participation in a cardio-pulmonary rehabilitation program based on exercise training within 8 weeks prior to screening • Any factor or condition likely to affect full participation in the study or compliance with the protocol (such as adherence to protocol mandated procedures), as judged by the investigator.
Availability of outcomes of interest			
WHO FC improvement	Yes (baseline to weeks 3, 6, 9, 12, 24, 36, and 48)	Yes (baseline to week 26)	Yes (baseline to week 24)
WHO FC worsening	Yes (baseline to weeks 3, 6, 9, 12, 24, 36, and 48)	Yes (baseline to week 26)	Yes (baseline to week 24)
Change in 6MWD	Yes (Hodges-Lehman median; baseline to weeks 3, 6, 9, 12, 24, 36, and 48)	Yes (Hodges-Lehman median; baseline to week 26)	Yes (least-squares mean treatment difference; baseline to week 24 or last postbaseline value)
Change in NT-proBNP	Yes (Hodges-Lehman median; baseline to weeks 3, 6, 9, 12, 24, 36, and 48)	Yes, Hodges-Lehman median; baseline to week 26	Yes (least-squares mean treatment difference; baseline to week 24 or last postbaseline value)
Time to death or non-fatal clinical worsening	Yes[a]	Yes[b]	No[c,d,e]
Time to death	Yes	Yes	No[c,d,e]
Time to PAH-related death	Yes	Yes	No[c,d,e]

[a] Time to death or nonfatal clinical worsening is defined as “death from any cause or specified nonfatal clinical worsening events (worsening-related listing for lung or heart–lung transplantation, initiation or rescue therapy with an approved background treatment or increase in the prostacyclin dose by ≥10%, atrial septostomy, hospitalization [≥24 hours] for worsening of pulmonary arterial hypertension, or worsening of pulmonary arterial hypertension relative to baseline as defined by both a worsened WHO functional class and a decrease in 6-minute walk distance by ≥15% [confirmed by two tests ≥4 hours but ≤1 week apart].”[9]

[b] Time to death or nonfatal clinical worsening is defined as a “composite of death or a complication related to pulmonary arterial hypertension, whichever occurred first, up to the end of the treatment period. Complications related to pulmonary arterial hypertension were disease progression or worsening of pulmonary arterial hypertension that resulted in hospitalization, initiation of parenteral prostanoid therapy or long-term oxygen therapy, or the need for lung transplantation or balloon atrial septostomy as judged by the physician. (Placement on a transplant waiting list represented an acute measure, as confirmed by the critical-event committee, and an actual lung transplantation would also meet this criterion.) Disease progression was defined as a decrease from baseline of at least 15% in the 6-minute walk distance (confirmed by means of a second test on a different day) accompanied by a worsening in WHO functional class (for the patients with WHO functional class II or III at baseline) or the need for additional treatment of pulmonary arterial hypertension (for the patients with WHO functional class III or IV at baseline).”

[c] Time to clinical worsening was addressed indirectly in the article when the authors acknowledge that [t]olerability for selexipag in TRACE was in line with that observed in GRIPHON and the known safety profile of selexipag.”

[d] This outcome is named “aggravation of PAH” and is defined as “death, transplantation, hospitalisation due to worsening PAH, or aggravation of PAH symptoms, i.e. a >=10% deterioration in 6-min walk distance or the need for additional PAH-specific therapies.”

[e] The authors remarked that there were no deaths in either group.

6MWD, 6-minute walking distance; NT-proBNP, N-terminal pro-brain natriuretic peptide; PAH, pulmonary arterial hypertension; PH, pulmonary hypertension; SoC, standard of care; WHO FC, World Health Organization functional class

Table 7 shows an overview of participant baseline characteristics for STELLAR and MAIC comparator studies (GRIPHON and TRACE).

Note that following an assessment of the study characteristics, several differences that may induce bias on estimated treatment effects were identified. While IPD exclusion and MAIC reweighting seek to improve trial comparability, some differences cannot be removed.

- Differences in SoC composition used in STELLAR vs comparator trials
- Different existing therapies, treatment guidelines and therapeutic pathways at the time of enrolment
- Differences in time since diagnosis with STELLAR having a longer period
- GRIPHON and TRACE contain no background triple therapy participants

Table 7: Overview of Baseline Patient Characteristics

Baseline Patient Characteristics	STELLAR (NCT04576988)		GRIPHON (NCT01106014)		TRACE (NCT03078907)	
	SoC (N=160)	sotatercept+SoC (N=163)	SoC (N=582)	Selexipag+SoC (N=574)	SoC (N=53)	Selexipag+SoC (N=55)
Age, year						
Mean (SD)	48.3 (15.5)	47.6 (14.1)	47.9 (15.55)	48.2 (15.19)	49.8 (13.63)	49 (14.75)
≥65	29 (18.1)	25 (15.3)	108 (18.6)	99 (17.2)	11 (20.8)	12 (21.8)
Female proportion, n (%)	127 (79.4)	129 (79.1)	466 (80.1)	457 (79.6)	42 (76.4)	35 (66.0)
Geographic region, n (%)						
Asia	5 (3.1%)	5 (3.1%)	113 (19.4)	115 (20.0)	0 (0)	0 (0)
Eastern Europe	14 (8.8%)	13 (8.0%)	155 (26.6)	149 (26.0)	0 (0)	0 (0)
Latin America	Not specified	Not specified	56 (9.6)	54 (9.4)	0 (0)	0 (0)
North America	51 (31.9%)	46 (28.2%)	98 (16.8)	95 (16.6)	8 (15.1)	6 (11.8)
Western Europe and Australia	56 (35.0%)	69 (42.3%)	160 (27.5)	161 (28.0)	42 (79.2)	46 (83.6)
Southern Europe	12 (7.5%)	11 (6.7%)	Not specified	Not specified	3 (5.6)	3 (5.5)
Central America	5 (3.1%)	3 (1.8%)	Not specified	Not specified	0 (0)	0 (0)
South America	15 (9.4%)	13 (8.0%)	Not specified	Not specified	0 (0)	0 (0)
PAH classification, n (%)						
Idiopathic PAH	106 (66.3)	83 (50.9)	337 (57.9)	312 (54.4)	42 (76.4)	40 (75.5)
Heritable PAH	24 (15.0)	35 (21.5)	13 (2.2)	13 (2.3)		
Connective tissue disease	19 (11.9)	29 (17.8)	167 (28.7)	167 (29.1)	10 (18.2)	8 (15.1)
Corrected-congenital shunts	7 (4.4)	9 (5.5)	50 (8.6)	60 (10.5)	1 (1.8)	4 (7.5)
HIV infection	0 (0)	0 (0)	5 (0.9)	5 (0.9)	1 (1.8)	0 (0)
Drug- and toxin-induced	4 (2.5)	7 (4.3)	10 (1.7)	17 (3.0)	1 (1.8)	1 (1.9)
WHO FC, n (%)						
I	0 (0)	0 (0)	5 (0.9)	4 (0.7)	0 (0)	0 (0)
II	78 (48.8)	79 (48.5)	255 (43.8)	274 (47.7)	41 (74.5)	33 (62.3)
III	82 (51.3)	84 (51.5)	314 (54.0)	293 (51.0)	14 (25.5)	20 (37.7)
IV	0 (0)	0 (0)	8 (1.4)	3 (0.5)	0 (0)	0 (0)
6MWD, meter, mean (SD)	404.7 (80.59)	397.6 (84.28)	348.0 (82.23)	358.5 (76.31)	449.5 (98.9)	453.1 (129.7)
Time from diagnosis to randomization, year, mean (SD)	8.28 (6.70)	9.22 (7.32)	2.5 (3.75)	2.3 (3.49)	Median: 2.8 Min: 0.4 Max: 22.1	Median: 3.2 Min: 0.4 Max: 16.5
SoC composition, n (%)						
None	0 (0)	0 (0)	124 (21.3)	112 (19.5)	0 (0)	0 (0)
ERA (monotherapy)	1 (0.6)	2 (1.2)	76 (13.1)	94 (16.4)	1 (1.8)	1 (1.9)
PDE-5i (monotherapy)	1 (0.6)	5 (3.1)	185 (31.8)	189 (32.9)	0 (0)	0 (0)
ERA+PDE-5i (double therapy)	38 (23.8)	36 (22.1)	197 (33.8)	179 (31.2)	Not specified	Not specified
ERA+PDE-5i or sGC stimulator (double therapy)	Not specified	Not specified	Not specified	Not specified	54 (98.2)	52 (98.1)
Other monotherapy[a]	2 (1.3)	2 (1.2)	0 (0)	0 (0)	0 (0)	0 (0)
Other double therapy[b]	17 (10.6)	20 (12.3)	0 (0)	0 (0)	0 (0)	0 (0)
Triple therapy[c]	101 (63.1)	98 (60.1)	0 (0)	0 (0)	0 (0)	0 (0)

[a] Including sGC monotherapy and PGI2 analogues monotherapy.

[b] Including ERA+sGC, ERA+prostacyclin, PDE-5i+prostacyclin, and sGC+prostacyclin.

[c] Including ERA+PDE-5i+prostacyclin and ERA+sGC+prostacyclin.

6MWD, 6-minute walking distance; ERA, endothelin receptor antagonist; HIV, human immunodeficiency virus; NT-proBNP, N-terminal pro-brain natriuretic peptide; PAH, pulmonary arterial hypertension; PDE-5i, phosphodiesterase type 5 inhibitor; PGI2, prostaglandin I2; PH, pulmonary hypertension; sGC, soluble guanylate cyclase; SD, standard deviation; SoC, standard of care; WHO FC, World Health Organization functional class

A.1.1.1 Effect Modifiers to be Adjusted in the Current MAIC Analysis

The following baseline characteristics, identified as potential effect modifiers based on previous PULSAR study versus GRIPHON comparisons, MSD Clinical Team input and expertise, and exploratory subgroup analyses of baseline characteristics, were selected as matching variables:

- Age [mean, SD]
    - For GRIPHON, matching on SD was done for Age measured in decades (10 years) to overcome numerical difficulties. The computed weights were then applied to the variables on the original scale (years).
- WHO Functional Class [Class II and Class III]
    - GRIPHON participants with WHO FC I and IV at baseline were excluded from matching on WHO FC to align with STELLAR.
- Diagnostic Group [Idiopathic/heritable, associated with CTD, Other (associated with corrected-congenital shunts, Drug/Toxin-induced, associated with HIV)]
    - Participants with PAH associated with HIV were only present in GRIPHON (10 participants), as HIV was an exclusion criterion in STELLAR.

For GRIPHON, matching on SD was done for 6MWD measured in km to overcome numerical difficulties. The computed weights were then applied to the variables on the original scale (m).

A1.2 MAIC Analysis

All trials sets include a common comparator (i.e. placebo). Therefore, an anchored MAIC between sotatercept and selexipag is conducted.

The inference procedure of MAIC follows two main steps:

- Matching step: re-weighting individual participants in the individual patient data (IPD) trial (i.e. STELLAR) to match the sample distribution of selected baseline characteristics reported in the aggregate data (AgD) trial (i.e. GRIPHON);
- Population-Adjusted inference of Treatment Effect: the adjusted treatment effect in the IPD trial (versus placebo) is derived using weighted statistics. Then, this adjusted treatment effect is compared to the treatment effect of the competitor’s intervention (versus placebo arm in competitor’s trial) in the AgD trial.

MAIC enables the estimation of treatment effect of sotatercept (with IPD) versus a competitor’s intervention selexipag (with AgD only) in the target population of the competitor’s trial, via adjusting the sample distribution of key baseline characteristics in the sotatercept trial by means of assigning individual weights.

A.1.2.1 Matching Step

MAIC moment-matching is accomplished by re-weighting individual participants from the trials with available IPD (i.e. STELLAR) to match the sample distribution of selected baseline characteristics reported for the trials with only aggregate data (i.e. GRIPHON).

The following data handling convention was adopted in the current MAIC analyses:

- when an individual has a missing value in any of the effect modifiers, its weight was also considered missing and hence not accounted for the subsequent calculation of the treatment effect of sotatercept versus selexipag;
- derived individual weights are rescaled to have a sum equal to the number of participants in IPD that have non-missing value in any of the effect modifiers.

In general, matching more effect modifiers and/or adjusting for bigger cross-trial difference may lead to extreme individual weights, which in turn diminish the effective sample size (ESS).

A.1.2.2 Population-Adjusted Inference of Treatment Effect

After the matching step, the adjusted treatment effect of sotatercept versus placebo is derived using weighted statistics based on IPD. The adjusted treatment effect of sotatercept versus placebo is then compared to the treatment effect of the competitor's intervention (selexipag versus placebo) in AgD, using Bucher method.

Importantly, excluding 198 participants from the IPD (STELLAR) due to more restrictive inclusion/exclusion criteria in the AgD trial (GRIPHON) may lead to imbalance of key baseline characteristics (e.g. effect modifiers and prognostic covariates), which may lead in turn to a biased estimate of within-trial treatment effect due to confounding. Since overall matching is applied in this MAIC, a post-hoc tabulation was produced to confirm that there were no impactful imbalances between the treatment arms (>10% difference in proportion in any category of the covariate) in IPD trial either before or after the matching step.

A.1.2.2.1. <u>WHO FC Endpoints</u>

A weighted Relative Risk (RR) between sotatercept and placebo is derived based on a generalized linear regression model with binomial response, log link, treatment as a covariate, and incorporating individual weights. Its 95% confidence interval (CI) is derived based on the sandwich (empirical) covariance matrix. Using the specified modelling approach enables weighted analysis, crucial for MAIC, as well as the use of robust sandwich estimator of variance. The latter is needed to account for uncertainty in weights, which are estimated and not fixed.

The treatment effect of selexipag versus placebo is also derived using a generalized linear regression model with binomial response, log link and treatment as a covariate, based on the risk per treatment arm of the selected endpoint reported in AgD. Of note, the reported treatment effect of selexipag in GRIPHON

study was quantified by odds ratio, based on a different method. Using the same statistical method to estimate RR further reduces the difference between the two studies and ensures a fair estimate of treatment effect of sotatercept relative to selexipag.

A2 RESULTS

There are a total of 323 participants in the FAS population of STELLAR. After excluding STELLAR participants who would not be eligible to enroll in GRIPHON (i.e. 198 STELLAR participants on triple background therapy for PAH), a total of 125 STELLAR participants were included in MAIC Population. Out of these, one participant in STELLAR MAIC population had a missing measurement of 6MWD at baseline, therefore was assigned a missing weight during MAIC and excluded from the after matching analysis results. This participant is counted in the results before matching and excluded in results after matching.

A2.1 GRIPHON

A.2.1.1 Effect Modifiers

A comparison of chosen effect modifiers before and after matching is presented in Table 8. The table shows that by means of MAIC participants from STELLAR MAIC population were matched to GRIPHON Study population. The highest differences between STELLAR MAIC population and GRIPHON Study population were observed in Heritable PAH and 6MWD at baseline and were accounted for after applying MAIC weights. This allows us to infer the potential sotatercept treatment effect on GRIPHON Study population and enables an adjusted indirect comparison between sotatercept and selexipag.

A.2.1.2 Participant and Disease Characteristics at Baseline

Participant and disease characteristics at baseline are presented in Table 9:. The following characteristics showed a mismatch of larger than 10% (difference of proportions or relative difference in means) between treatment groups after matching:

- Ethnicity Hispanic (more in placebo arm),
- Idiopathic PAH (more in placebo arm) – similar differences observed in the STELLAR FAS population,
- WHO FC II (more in placebo arm).

Although sotatercept suggested to provide treatment effect across all prespecified subgroups in STELLAR, a greater proportion of WHO FC II participants in placebo arm could potentially favor placebo in within-STELLAR treatment effect estimate and consequently bias ITC estimate in favor of selexipag.

A.2.1.3 Treatment Effect Analyses

A.2.1.3.1. WHO FC Endpoints

The MAIC analysis results for WHO FC improvement, maintenance and worsening are presented in Table 10:, Table 11, and Table 12, respectively.

A3 TABLES AND FIGURES

A3.1 GRIPHON

A.3.1.1 Effect Modifiers

Table 8: Effect Modifiers - Matching-Adjusted Indirect Comparison of sotatercept versus Selexipag

	GRIPHON[a]	Before Matching STELLAR (MK-7962-003)[b]	After Matching STELLAR (MK-7962-003)[b]
	(N=1156)	(N=125)	
Age (Years)			
Mean (SD)	48.0 (15.4)	50.6 (15.8)	
WHO Functional Class[d], %			
Class II	46.6	48.8	
Class III	53.4	51.2	
WHO Diagnostic Group, %			
Idiopathic PAH	56.1	56	
Heritable PAH	2.2	13.6	
Other[e]	12.7	8.8	
PAH Associated with Connective Tissue Disease	28.9	21.6	
Baseline 6MWD (m)			
Mean (SD)	353.2 (79.9)	384.4 (90.1)	

a: GRIPHON: Sitbon et al., (2015): Selexipag for the Treatment of Pulmonary Arterial Hypertension. N Engl J Med. 373(26):2522-33. doi: 10.1056/NEJMoa1503184. PMID: 26699168

b: Number of participants: Full Analysis Set population in STELLAR, excluding participants on triple background therapy for PAH; Database Cutoff Date: 06DEC2022

c: Effective sample size computed as the square of the summed weights divided by the sum of the squared weights

d: 20 participants in GRIPHON with Class I or IV at baseline were excluded, percentages and total N were recalculated based on Class II and III participants only. Matching on this variable was based on recalculated percentages

e: Includes drug-induced or toxin induced PAH, PAH associated with corrected congenital shunts, and PAH associated with HIV (10 participants in GRIPHON only, as HIV was an exclusion criterion in STELLAR)

6MWD: 6-Minute Walk Distance; HIV: Human Immunodeficiency Virus; PAH: Pulmonary Arterial Hypertension; SD: Standard Deviation; WHO: World Health Organization

A.3.1.2 Participant and Disease Characteristics at Baseline

Table 9: Participant and Disease Characteristics at Baseline in STELLAR - Matching Adjusted Indirect Comparison of sotatercept versus Selexipag - (STELLAR MAIC Population Matched to GRIPHON Study Population)

Study: STELLAR (MK-7962-003)[a]	Before Matching[b]			After Matching[c,d]		
	sotatercept	Placebo	Total	sotatercept	Placebo	Total
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Participants in population	65	60	125			
Age (Years)						
Participants with data	65	60	125			
Mean	49.7	51.5	50.6			
SD	15.80	16.06	15.88			
Median	52.0	51.0	51.0			
IQR	37.0,60.0	39.5,64.5	38.0,63.0			

Study: STELLAR (MK-7962-003)[a]	Before Matching[b]			After Matching[c,d]		
	sotatercept	Placebo	Total	sotatercept	Placebo	Total
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Range	18.0 to 81.0	18.0 to 82.0	18.0 to 82.0			
Age Group (Years)						
18 to <45	24 (36.92)	20 (33.33)	44 (35.20)			
45 to <65	29 (44.62)	25 (41.67)	54 (43.20)			
65 to <75	8 (12.31)	11 (18.33)	19 (15.20)			
≥75	4 (6.15)	4 (6.67)	8 (6.40)			
Sex						
Male	13 (20.00)	13 (21.67)	26 (20.80)			
Female	52 (80.00)	47 (78.33)	99 (79.20)			
Race						
White	57 (87.69)	53 (88.33)	110 (88.00)			
Black or African American	0 (0.00)	2 (3.33)	2 (1.60)			
Asian	0 (0.00)	2 (3.33)	2 (1.60)			
Multiple	1 (1.54)	0 (0.00)	1 (0.80)			
Other	4 (6.15)	2 (3.33)	6 (4.80)			
Missing	3 (4.62)	1 (1.67)	4 (3.20)			

Please return to: **NICE DOCS**

Study: STELLAR (MK-7962-003)[a]	Before Matching[b]			After Matching[c,d]		
	sotatercept	Placebo	Total	sotatercept	Placebo	Total
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Ethnicity						
Hispanic or Latino	13 (20.00)	18 (30.00)	31 (24.80)			
Not Hispanic or Latino	49 (75.38)	39 (65.00)	88 (70.40)			
Not Reported	3 (4.62)	3 (5.00)	6 (4.80)			
Height (cm)						
Participants with data	65	60	125			
Mean	164.9	165.1	165.0			
SD	8.12	8.86	8.45			
Median	165.0	164.5	165.0			
IQR	160.0,168.0	158.5,170.6	160.0,169.0			
Range	144.0 to 193.0	147.3 to 184.0	144.0 to 193.0			
Weight (kg)						
Participants with data	65	60	125			
Mean	69.9	73.2	71.5			
SD	17.32	17.02	17.19			
Median	67.0	71.8	68.2			
IQR	57.2,78.0	62.0,82.2	60.0,80.2			

Study: STELLAR (MK-7962-003)[a]	Before Matching[b]			After Matching[c,d]		
	sotatercept	Placebo	Total	sotatercept	Placebo	Total
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Range	42.9 to 135.0	46.6 to 141.3	42.9 to 141.3			
BMI (kg/m²)						
Participants with data	65	60	125			
Mean	25.5	26.9	26.2			
SD	5.03	6.18	5.63			
Median	24.7	25.5	25.3			
IQR	21.9,28.0	22.9,29.5	22.6,29.0			
Range	16.6 to 39.0	18.4 to 54.2	16.6 to 54.2			
Time Since PAH Diagnosis (Years)						
Participants with data	65	60	125			
Mean	9.5	8.2	8.9			
SD	9.05	7.91	8.51			
Median	5.8	5.4	5.7			
IQR	2.8,13.4	2.2,12.2	2.4,12.4			
Range	0.7 to 38.3	0.6 to 40.2	0.6 to 40.2			
WHO Diagnostic Group						
Idiopathic PAH	34 (52.31)	36 (60.00)	70 (56.00)			

Study: STELLAR (MK-7962-003)[a]	Before Matching[b]			After Matching[c,d]		
	sotatercept	Placebo	Total	sotatercept	Placebo	Total
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Heritable PAH	9 (13.85)	8 (13.33)	17 (13.60)			
Other[e]	8 (12.31)	3 (5.00)	11 (8.80)			
PAH Associated with Connective Tissue Disease	14 (21.54)	13 (21.67)	27 (21.60)			
WHO Functional Class						
Class II	31 (47.69)	30 (50.00)	61 (48.80)			
Class III	34 (52.31)	30 (50.00)	64 (51.20)			
Type of Background PAH Therapy						
Mono	9 (13.85)	4 (6.67)	13 (10.40)			
Double	56 (86.15)	56 (93.33)	112 (89.60)			
Exercise Program for Cardiopulmonary Rehabilitation						
Yes	6 (9.23)	5 (8.33)	11 (8.80)			
No	59 (90.77)	55 (91.67)	114 (91.20)			
Reason for not Being on Exercise Program						
Not reimbursable	6 (9.23)	13 (21.67)	19 (15.20)			
Not recommended for this patient	25 (38.46)	20 (33.33)	45 (36.00)			
Not available locally	17 (26.15)	16 (26.67)	33 (26.40)			
Refused by the patient	11 (16.92)	6 (10.00)	17 (13.60)			

Study: STELLAR (MK-7962-003)[a]	Before Matching[b]			After Matching[c,d]		
	sotatercept	Placebo	Total	sotatercept	Placebo	Total
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Missing	6 (9.23)	5 (8.33)	11 (8.80)			
Hemoglobin (g/dl)						
Participants with data	64	60	124			
Mean	139.3	134.7	137.1			
SD	15.69	16.56	16.22			
Median	141.0	135.7	139.0			
IQR	131.5,150.5	125.5,149.0	126.5,150.0			
Range	94.0 to 166.0	100.0 to 163.0	94.0 to 166.0			
Systolic Blood Pressure (mmHg)						
Participants with data	65	60	125			
Mean	118.2	115.1	116.7			
SD	15.36	17.25	16.30			
Median	118.0	112.5	115.0			
IQR	105.0,130.0	100.0,130.0	104.0,130.0			
Range	90.0 to 152.0	86.0 to 154.0	86.0 to 154.0			
Diastolic Blood Pressure (mmHg)						
Participants with data	65	60	125			

Study: STELLAR (MK-7962-003)[a]	Before Matching[b]			After Matching[c,d]		
	sotatercept	Placebo	Total	sotatercept	Placebo	Total
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Mean	70.0	70.0	70.0			
SD	12.56	10.03	11.37			
Median	67.0	70.0	68.0			
IQR	61.0,80.0	62.0,77.0	62.0,78.0			
Range	44.0 to 111.0	40.0 to 93.0	40.0 to 111.0			
Pulse Rate (beat/min)						
Participants with data	65	60	125			
Mean	79.0	75.0	77.1			
SD	12.94	12.14	12.67			
Median	77.0	75.5	76.0			
IQR	70.0,87.0	66.5,82.0	69.0,85.0			
Range	47.0 to 110.0	50.0 to 120.0	47.0 to 120.0			
6-Minute Walk Distance (m)						
Participants with data	65	59	124			
Mean	378.6	390.7	384.4			
SD	93.19	87.66	90.44			

Study: STELLAR (MK-7962-003)[a]	Before Matching[b]			After Matching[c,d]		
	sotatercept	Placebo	Total	sotatercept	Placebo	Total
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Median	403.5	414.0	407.3			
IQR	318.5,461.0	322.0,465.0	321.8,462.5			
Range	160.5 to 497.5	157.0 to 494.5	157.0 to 497.5			
Cardiac Index (L/min/m²)						
Participants with data	65	60	125			
Mean	2.6	2.7	2.7			
SD	0.59	0.69	0.64			
Median	2.5	2.6	2.6			
IQR	2.2,3.0	2.2,3.1	2.2,3.0			
Range	1.0 to 4.2	1.4 to 4.9	1.0 to 4.9			
Pulmonary Vascular Resistance (dynes*sec/cm⁵)						
Participants with data	65	60	125			
Mean	770.0	660.1	717.2			
SD	382.57	236.90	324.47			
Median	656.0	600.0	624.0			
IQR	512.0,856.0	480.0,760.0	496.0,800.0			

Study: STELLAR (MK-7962-003)[a]	Before Matching[b]			After Matching[c,d]		
	sotatercept	Placebo	Total	sotatercept	Placebo	Total
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Range	400.0 to 2200.0	400.0 to 1496.0	400.0 to 2200.0	████	████	████

a: Database Cutoff Date: 06DEC2022

b: Number of participants: STELLAR MAIC population (Full Analysis Set, excluding participants on triple background therapy for PAH)

c: Number of participants: sum of rescaled weights of participants from STELLAR MAIC population with non-missing values in all effect modifiers

d: MAIC of STELLAR (Database Cutoff Date: 06DEC2022) with GRIPHON Study, adjusted for the following effect modifiers: Age, WHO Functional Class, WHO Diagnostic Group, 6-Minute Walk Distance

e: Includes drug-induced or toxin induced PAH, PAH associated with corrected congenital shunts, and PAH associated with HIV (10 participants in GRIPHON only, as HIV was an exclusion criterion in STELLAR)

BMI: Body Mass Index; IQR: Interquartile Range; MAIC: Matching-Adjusted Indirect Comparison; PAH: Pulmonary Arterial Hypertension; SD: Standard Deviation; WHO: World Health Organization

A.3.1.3 Treatment Effect Analyses

A.3.1.3.1. WHO FC Endpoints

A.3.1.3.1.1. WHO FC Improvement

Table 10: Analysis of Improvement from Baseline in WHO Functional Classification at Week 24 (STELLAR) / Week 26 (GRIPHON) - Based on Complete Cases - Matching Adjusted Indirect Comparison Analysis of - sotatercept versus Selexipag - (GRIPHON Study Population and STELLAR MAIC Population)

	Before Matching					After Matching[a]				
Study Treatment Arm	**N[b]**	**Participants with Event n (%)**	**Risk Ratio[c] [95%-CI]**	**ITC Risk Ratio[d] [95%-CI]**	**p-Value[e]**	**N[f]**	**Participants with Event n[f] (%)**	**Risk Ratio[g] [95%-CI][h]**	**ITC Risk Ratio[d] [95%-CI]**	**p-Value[e]**
STELLAR (MK-7962-003)[i]										
sotatercept	62	17 (27.4)	1.48 [0.74; 2.95]							
Placebo	54	10 (18.5)								
GRIPHON[j]										
Selexipag	471	75 (15.9)	1.51 [1.08; 2.12]							
Placebo	456	48 (10.5)								

a: MAIC of STELLAR (Database Cutoff Date: 06DEC2022) with GRIPHON Study, adjusted for the following effect modifiers: Age, WHO Functional Class, WHO Diagnostic Group, 6-Minute Walk Distance

b: Number of participants: STELLAR MAIC population (Full Analysis Set, excluding participants on triple background therapy for PAH), and GRIPHON Study population (all randomized participants). Participants with WHO FC II or III at Baseline and non-missing WHO FC at Week 24/Week 26

c: Based on a generalized linear model with binomial response, log link and treatment as a covariate

Please return to: **NICE DOCS**

d: Bucher methodology using separate study results (estimate and its standard error) with a common control arm (Placebo) to perform indirect comparison of sotatercept (STELLAR) versus Selexipag (GRIPHON)
e: Two-sided p-value calculated from the test statistic associated with the ITC estimate and its standard error
f: Number of participants: sum of rescaled weights of participants from STELLAR MAIC population with non-missing values in all effect modifiers
g: For STELLAR, based on a generalized linear model with binomial response, log link, treatment as a covariate and incorporating individual weights. For GRIPHON, based on the same method as before matching
h: Based on a robust sandwich estimator using PROC GENMOD in SAS
i: Database Cutoff Date: 06DEC2022
j: GRIPHON: Sitbon et al., (2015): Selexipag for the Treatment of Pulmonary Arterial Hypertension. N Engl J Med. 373(26):2522-33. doi: 10.1056/NEJMoa1503184. PMID: 26699168
CI: Confidence Interval; PAH: Pulmonary Arterial Hypertension; WHO: World Health Organization

A.3.1.3.1.2. WHO FC Maintenance

Table 11: Analysis of Maintenance from Baseline in WHO Functional Classification at Week 24 (STELLAR) / Week 26 (GRIPHON) - Based on Complete Cases - Matching Adjusted Indirect Comparison Analysis of - sotatercept versus Selexipag - (GRIPHON Study Population and STELLAR MAIC Population)

	Before Matching					**After Matching[a]**				
Study Treatment Arm	**N[b]**	**Participants with Event n (%)**	**Risk Ratio[c] [95%-CI]**	**ITC Risk Ratio[d] [95%-CI]**	**p-Value[e]**	**N[f]**	**Participants with Event n[f] (%)**	**Risk Ratio[g] [95%-CI][h]**	**ITC Risk Ratio[d] [95%-CI]**	**p-Value[e]**
STELLAR (MK-7962-003)[i]										
sotatercept	62	43 (69.4)	1.01 [0.79; 1.29]			███	███	███	███	███
Placebo	54	37 (68.5)		███	███	███	███	███	███	███

GRIPHON[j]										
Selexipag	471	365 (77.5)	0.93 [0.88; 1.00]							
Placebo	456	378 (82.9)								

a: MAIC of STELLAR (Database Cutoff Date: 06DEC2022) with GRIPHON Study, adjusted for the following effect modifiers: Age, WHO Functional Class, WHO Diagnostic Group, 6-Minute Walk Distance

b: Number of participants: STELLAR MAIC population (Full Analysis Set, excluding participants on triple background therapy for PAH), and GRIPHON Study population (all randomized participants). Participants with WHO FC II or III at Baseline and non-missing WHO FC at Week 24/Week 26

c: Based on a generalized linear model with binomial response, log link and treatment as a covariate

d: Bucher methodology using separate study results (estimate and its standard error) with a common control arm (Placebo) to perform indirect comparison of sotatercept (STELLAR) versus Selexipag (GRIPHON)

e: Two-sided p-value calculated from the test statistic associated with the ITC estimate and its standard error

f: Number of participants: sum of rescaled weights of participants from STELLAR MAIC population with non-missing values in all effect modifiers

g: For STELLAR, based on a generalized linear model with binomial response, log link, treatment as a covariate and incorporating individual weights. For GRIPHON, based on the same method as before matching

h: Based on a robust sandwich estimator using PROC GENMOD in SAS

i: Database Cutoff Date: 06DEC2022

j: GRIPHON: Sitbon et al., (2015): Selexipag for the Treatment of Pulmonary Arterial Hypertension. N Engl J Med. 373(26):2522-33. doi: 10.1056/NEJMoa1503184. PMID: 26699168

CI: Confidence Interval; PAH: Pulmonary Arterial Hypertension; WHO: World Health Organization

A.3.1.3.1.3. WHO FC Worsening

Table 12: Analysis of Worsening from Baseline in WHO Functional Classification at Week 24 (STELLAR) / Week 26 (GRIPHON) - Based on Complete Cases - Matching Adjusted Indirect Comparison Analysis of - sotatercept versus Selexipag – (GRIPHON Study Population and STELLAR MAIC Population)

	Before Matching					After Matching[a]				
Study Treatment Arm	N[b]	Participants with Event n (%)	Risk Ratio[c] [95%-CI]	ITC Risk Ratio[d] [95%-CI]	p-Value[e]	N[f]	Participants with Event n[f] (%)	Risk Ratio[g] [95%-CI][h]	ITC Risk Ratio[d] [95%-CI]	p-Value[e]
STELLAR (MK-7962-003)[i]										
sotatercept	62	2 (3.2)	0.25 [0.05; 1.15]							
Placebo	54	7 (13.0)								
GRIPHON[j]										
Selexipag	471	31 (6.6)	1.00 [0.62; 1.63]							
Placebo	456	30 (6.6)								

a: MAIC of STELLAR (Database Cutoff Date: 06DEC2022) with GRIPHON Study, adjusted for the following effect modifiers: Age, WHO Functional Class, WHO Diagnostic Group, 6-Minute Walk Distance

b: Number of participants: STELLAR MAIC population (Full Analysis Set, excluding participants on triple background therapy for PAH), and GRIPHON Study population (all randomized participants). Participants with WHO FC II or III at Baseline and non-missing WHO FC at Week 24/Week 26

c: Based on a generalized linear model with binomial response, log link and treatment as a covariate

d: Bucher methodology using separate study results (estimate and its standard error) with a common control arm (Placebo) to perform indirect comparison of sotatercept (STELLAR) versus Selexipag (GRIPHON)

e: Two-sided p-value calculated from the test statistic associated with the ITC estimate and its standard error

f: Number of participants: sum of rescaled weights of participants from STELLAR MAIC population with non-missing values in all effect modifiers

g: For STELLAR, based on a generalized linear model with binomial response, log link, treatment as a covariate and incorporating individual weights. For GRIPHON, based on the same method as before matching

h: Based on a robust sandwich estimator using PROC GENMOD in SAS

i: Database Cutoff Date: 06DEC2022

j: GRIPHON: Sitbon et al., (2015): Selexipag for the Treatment of Pulmonary Arterial Hypertension. N Engl J Med. 373(26):2522-33. doi: 10.1056/NEJMoa1503184. PMID: 26699168

CI: Confidence Interval; PAH: Pulmonary Arterial Hypertension; WHO: World Health Organization

Appendix B Revised Company base case results

B1 Model parameters

A summary of model inputs and parameters informing the updated company base case is presented in Table 13 below.

Table 13. Table of updated company base case model parameters and measurements of uncertainty

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)
Model general settings		
Perspective	NHS and PSS	N/A
Time horizon	Lifetime (30 years)	N/A
Discount rate: costs and outcomes	3.5%	N/A
Model structure	Risk-based	N/A
Patient characteristics		
Mean age, years	47.9	0.8 (Normal)
Females, %	79.3%	2.3% (Beta)
Body weight, kg	███	███████
Age-specific mortality rate	UK national life tables rates by age and sex	N/A
Markov Model		
Initial distribution	All patients in intermediate-low risk on sotatercept/ selexipag + dual therapy	N/A
Week-24 distribution (sotatercept + dual therapy) from intermediate-low risk	Low Risk: ███ Intermediate-low risk: ███ Intermediate-high risk tunnel state: 4.3% High risk: 2.9%	(Dirichlet) Low risk: █████ Intermediate-low risk: █████ Intermediate-high risk tunnel state:: 2.5██ High risk: ████
Week-24 distribution (selexipag + dual therapy) from intermediate-low risk	Low Risk: ███ Intermediate-low risk: ███ Intermediate-high risk: ███ High risk history of or on PCA: █	(Dirichlet) Low risk: █████ Intermediate-low risk: █████ Intermediate-high risk: █████ High risk history of or on PCA: ███
Long-term transition probabilities (sotatercept + dual therapy)	See Table 82 in company submission and updated transition probabilities for new health states in the response above.	NA
RR, RS worsening, selexipag vs. sota	████████████████	████████████████████
RR, RS improvement, selexipag vs. sota	████████████████	████████████████████

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)
Baseline age for mortality adjustment	51.7	10.3 (Normal)
Mortality distribution for single parametric model fitted to low risk	Gamma	NA
HR for mortality (relative to low risk)	Intermediate-low risk (including discon.): ████ Intermediate-high risk (including discon.): ████ High risk (including discon.): ████	(Log-normal) Intermediate-low risk: ████ Intermediate-high risk: ████ High risk: ████
Post-transplant 12-week mortality risk	2.8%	0.6% (Beta)
PAH hospitalisation risk	Low risk (including discon.): 1.8% Intermediate-low risk (including discon.): 2.7% Intermediate-high risk (including discon.): 4.3% High risk (including discon.): 5.7% Lung/heart transplant:1.8%	(Beta) Low risk: 0.3% Intermediate-low risk: 0.2% Intermediate-high risk: 0.2% High risk: 0.6% Lung/heart transplant: 0.3%
HR for PAH hospitalisation, sotatercept vs. selexipag in low and intermediate-low risk health states	████	████ (Log-normal)
Lung/heart transplant risk	Intermediate-high risk: 0.05% High risk: 0.07%	0.01% (Beta)
Utility		
Utility by health state	Low risk (including discon.): ████ Intermediate-low risk (including discon.): ████ Intermediate-high risk (including discon.): ████ High risk (including discon.): ████ Heart/Lung Transplant: ████	(Beta) Low risk: ████ Intermediate-low risk: ████ Intermediate-high risk: ████ High risk: ████ Heart/lung transplant: ████
Disutility for treatment administration	IV infusion: 0.307 SC injection: 0 Inhaled: 0 Oral: 0	(Normal) IV infusion: 0.061 SC injection: 0 Inhaled: 0 Oral: 0
Hospitalisation disutility	0.071	0.014 (Normal)
Transplant disutility	0.071	0.014 (Normal)
Carer utility gain	Low risk: -0.036 Intermediate-low risk: -0.023 Intermediate-high risk: -0.013 High risk:0	(Normal) Low risk: -0.007 Intermediate-low risk: -0.005 Intermediate-high risk: -0.003 High risk:0

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)
Costs		
Drug acquisition costs (12-weeks) Branded drug costs were not included in the one-way and probabilistic sensitivity analyses, given the precision of the input value.	Bosentan, oral: £126.27 Ambrisentan, oral: £96.85 Sildenafil citrate, oral: £114.21 Tadalafil, oral: £55.16 Treprostinil, IV infusion: £3,570.70 Iloprost, inhaled: £9244.46	(Gamma) Bosentan, oral: £25.25 Ambrisentan, oral: £19.37 Sildenafil citrate, oral: £22.84 Tadalafil, oral: £11.03 Treprostinil, IV infusion: £714.14 Iloprost, inhaled: £1848.89
Number of administrations in 12 weeks	Sotatercept, SC injection: 4 Bosentan, oral:168 Ambrisentan, oral: 84 Macitentan, oral: 84 Sildenafil citrate, oral: 252 Tadalafil, oral: 84 Epoprostenol, IV infusion: 84 Treprostinil, IV infusion: NA Iloprost, inhaled: 630 Selexipag, oral: 168	NA
Treatment distribution sotatercept (low risk and intermediate-low risk)	Sotatercept, SC injection: 100% Bosentan, oral:12.8% Ambrisentan, oral: 43.3% Macitentan, oral: 44.0% Sildenafil citrate, oral: 61.9% Tadalafil, oral: 38.1% Epoprostenol, IV infusion: 0% Treprostinil, IV infusion: 0% Iloprost, inhaled: 0% Selexipag, oral: 0%	(Beta) Sotatercept, SC injection: 20% Bosentan, oral:3% Ambrisentan, oral: 8% Macitentan, oral: 9% Sildenafil citrate, oral: 13% Tadalafil, oral: 7% Epoprostenol, IV infusion: 0% Treprostinil, IV infusion: 0% Iloprost, inhaled: 0% Selexipag, oral: 0%
Treatment distribution sotatercept (intermediate-high risk tunnel state, intermediate-low risk history of or on PCA, intermediate-high risk history of or on PCA)	Scenario 1: Sotatercept, SC injection: 100% Bosentan, oral:12.8% Ambrisentan, oral: 43.3% Macitentan, oral: 44.0% Sildenafil citrate, oral: 61.9% Tadalafil, oral: 38.1% Epoprostenol, IV infusion: 61.9% Treprostinil, IV infusion: 5.0% Iloprost, inhaled: 33.1% Selexipag, oral: 0% Scenario 2:	(Beta) Sotatercept, SC injection: 0% Bosentan, oral:3% Ambrisentan, oral: 8% Macitentan, oral: 9% Sildenafil citrate, oral: 13% Tadalafil, oral: 7% Epoprostenol, IV infusion: NA Treprostinil, IV infusion: NA Iloprost, Inhaled: NA Selexipag, oral: 0% The treatment utilisation of epoprostenol, Treprostinil and iloprost are not varied in the

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)
	Sotatercept, SC injection: 100% Bosentan, oral:12.8% Ambrisentan, oral: 43.3% Macitentan, oral: 44.0% Sildenafil citrate, oral: 61.9% Tadalafil, oral: 38.1% Epoprostenol, IV infusion: 61.9% Treprostinil, IV infusion: 5.0% Iloprost, inhaled: 33.1% Selexipag, oral: 0%	DSA or PSA in the updated company economic model. The treatment utilisation of PCA is accounted for through the weighted approach and therefore would not be feasible or appropriate to vary these in the sensitivity analysis.
Treatment distribution sotatercept (high risk history of or on PCA)	Sotatercept, SC injection: 15% Bosentan, oral:12.8% Ambrisentan, oral: 43.3% Macitentan, oral: 44.0% Sildenafil citrate, oral: 61.9% Tadalafil, oral: 38.1% Epoprostenol, IV infusion: 52.7% Treprostinil, IV infusion: 4.2% Iloprost, inhaled: 18.1% Selexipag, oral: 0%	(Beta) Sotatercept, SC injection: 0% Bosentan, oral:3% Ambrisentan, oral: 8% Macitentan, oral: 9% Sildenafil citrate, oral: 13% Tadalafil, oral: 7% Epoprostenol, IV infusion: 12% Treprostinil, IV infusion: 0% Treprostinil, SC infusion: 0% Iloprost, Inhaled: 7% Selexipag, oral: 0%
Treatment distribution selexipag (low risk and intermediate-low risk)	Sotatercept, SC injection: 0% Bosentan, oral: 12.8% Ambrisentan, oral: 43.3% Macitentan, oral: 44.0% Sildenafil citrate, Oral: 61.9% Tadalafil, oral: 38.1% Epoprostenol, IV infusion: 0% Treprostinil, IV infusion: 0% Iloprost, inhaled: 0% Selexipag, oral: 100%	(Beta) Sotatercept, SC injection: 0% Bosentan, oral:3% Ambrisentan, oral: 8% Macitentan, oral: 9% Sildenafil citrate, oral: 13% Tadalafil, oral: 7% Epoprostenol, IV infusion: 0% Treprostinil, IV infusion: 0% Iloprost, inhaled: 0% Selexipag, oral: 20%
Treatment distribution selexipag (intermediate-high risk tunnel state, intermediate-low history of or on PCA, intermediate-high history of or on PCA and high risk history of or on PCA)	Sotatercept, SC injection: 0% Bosentan, oral:12.8% Ambrisentan, oral: 43.3% Macitentan, oral: 44.0% Sildenafil citrate, oral: 61.9% Tadalafil, oral: 38.1% Epoprostenol, IV infusion: 52.7% Treprostinil, IV infusion: 4.2% Iloprost, inhaled: 18.1% Selexipag, oral: 15%	(Beta) Sotatercept, SC injection: 0% Bosentan, oral:3% Ambrisentan, oral: 8% Macitentan, oral: 9% Sildenafil citrate, oral: 13% Tadalafil, oral: 7% Epoprostenol, IV infusion: 12% Treprostinil, IV infusion: 0% Iloprost, inhaled: 7% Selexipag, oral: 0%

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)
Treatment distribution selexipag (intermediate-low history on or off PCA)	Sotatercept, SC injection: 0% Bosentan, oral:12.8% Ambrisentan, oral: 43.3% Macitentan, oral: 44.0% Sildenafil citrate, oral: 61.9% Tadalafil, oral: 38.1% Epoprostenol, IV infusion: 61.9% Treprostinil, IV infusion: 5.0% Iloprost, inhaled: 33.1% Selexipag, oral: 0%	(Beta) Sotatercept, SC injection: 0% Bosentan, oral:3% Ambrisentan, oral: 8% Macitentan, oral: 9% Sildenafil citrate, oral: 13% Tadalafil, oral: 7% Epoprostenol, IV infusion: 12% Treprostinil, IV infusion: 0% Iloprost, inhaled: 7% Selexipag, oral: 0%
One off costs: sotatercept admin costs with average of 2.5 appointments per patient	One off costs: £1,110.00	One off costs: £222.000 (Gamma)
One-Off HCRU: Sotatercept pre-admin costs	Cost per admin: £168.00 Average number of admins: 3.5	NA
One-Off HCRU: Sotatercept Blood Tests	Cost per Test: £20 Number of Tests: 5	(Gamma) £4.00 1.00
Recurring HCRU costs by health state per 12-week cycle (sotatercept + dual therapy)	Low risk: £0 Intermediate-low risk: £0 Intermediate-high risk: £0 High risk: £2,555	(Gamma) £0 £0 £0 £511
Recurring HCRU costs by health state (selexipag + dual therapy)	Low risk: £302 Intermediate-low risk: £302 Intermediate-high risk: £1,966 High risk: £1,966	£60 £60 £393 £393
Recurring HCRU costs by health state (discontinued sotatercept + dual therapy)	Low risk: £0 Intermediate-low risk: £0 Intermediate-high risk: £1,966 High risk: £2,555	£0 £0 £393 £511
Cost of hospitalisation event	£1,825.14	£365.03 (Gamma)
Cost of transplant event	£78,315.68	£15,663.14 (Gamma)
Mortality event cost	£15,743.00	£3,148.60 (Gamma)

B2 Deterministic model results

The revised company base case results, for initiating sotatercept + dual background therapy vs. selexipag + dual background therapy in intermediate-low risk patients, incorporating the updated settings specified in the responses above are presented in Table 14.

Sotatercept in combination with dual therapy accumulated ████ of costs and ██ QALYs and selexipag in combination with dual therapy accumulated ████ of costs and ██ QALYs. The resulting ICER is ████ per QALY, inclusive of the 1.2 QALY severity modifier for which sotatercept qualifies.

Table 14: Deterministic base case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Sotatercept + dual therapy	████	██	██	████	██	4.09	████	████
Selexipag + dual therapy	████	██	██	-	-	-	-	-

ICER: incremental cost-effectiveness ratio; LYG: life-years gained; QALY: quality-adjusted life-years

Table 15 presents the incremental deterministic net health benefit (NHB) and net monetary benefit (NMB) versus sotatercept.

Table 15: Deterministic net health benefit and net monetary benefit

Technologies	Incremental costs (£)	Incremental QALYs	NHB at £20,000	NHB at £30,000	NMB at £20,000	NMB at £30,000
Selexipag + dual therapy	████	4.09* 1.2 = 4.91	██	██	████	████

NHB: net health benefit; NMB: net monetary benefit; LYG: life-years gained; QALY: quality-adjusted life-years

B3 Exploring uncertainty

B3.1 Deterministic sensitivity analysis

A one-way deterministic sensitivity analysis was conducted to identify model parameters that have the greatest independent impact on the model results. A tornado diagram including the top ten most influential parameters is presented in Figure 4, below.

Figure 4: Tornado diagram

Abbreviations: BGT, background therapy; HCRU, healthcare resource use; IH, intermediate-high; IL, intermediate-low; IV, intravenous; PAH, pulmonary arterial hypertension; QALY, quality-adjusted life-years; RR, relative risk; RS, risk status

B3.2 Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was run for 1,000 iterations. A weighted probabilistic ICER is generated using the proportion of iterations that met the 1.2 severity modifier threshold.

The parameters varied in the PSA and the nature of the uncertainty around their value are described in

Table 13 above.

The mean summary results from the probabilistic analysis are presented in Table 16 below. Graphical results including the cost-effectiveness plane (incremental costs vs. incremental QALYs) and the cost-effectiveness acceptability curve (CEAC) are shown below in Figure 5 and Figure 6 respectively. Sotatercept was demonstrated a probability of being a cost-effective treatment initiated in intermediate-low risk patients compared to selexipag in ██ and ██ of iterations at a willingness-to-pay threshold of £20,000/QALY and £30,000/QALY, respectively.

Table 16: Probabilistic base case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	ICER with severity modifier (£/QALY)	ICER weighted according to modifier (£/QALY)
Severity modifier							████	████	████
Sotatercept + dual therapy	████	█	██	████	█	3.75	████	████	████
Selexipag + dual therapy	████	█	██	-	-	-	-	-	-

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

Figure 5: Cost-effectiveness plane scatter plot (1,000 iterations) – Incremental Costs and QALYs (severity modifier applied)

█

Abbreviations: QALY, quality-adjusted life-year; WTP, willingness-to-pay

Figure 6: Cost-effectiveness acceptability curve (1,000 iterations) (severity modifier applied)

█

Abbreviations: BGT, background therapy; QALY, quality-adjusted life-year

B3.3 Scenario analysis results

A number of additional scenario were explored to understand the impact of structural and input assumptions in the model base case that have been discussed in this response document. The list of scenarios and the rationale behind their inclusion in the analysis is presented in Table 17.

In addition, the full set of scenario analyses are replicated with the updated company base case but including drug wastage costs for sotatercept. Both sets of scenario analysis results without and with sotatercept drug wastage costs are presented in Table 18 and Table 19, respectively.

Table 17: List of scenario analysis

Scenario	Rationale
Mortality approach: Dependent model using Gompertz	The dependent model, jointly fitting a parametric model to each risk status Kaplan-Meier curve from Rosenkranz, to estimate long-term overall survival is tested in scenarios. The evolution of risks over time, as provided by the Gompertz distribution, was judged to be the most appropriate alternatives to the base case choice and therefore included in the scenario analyses.
Selexipag long-term transitions: within-trial analysis, no RR reduction	Acknowledging limitations associated with both the MAIC and WTA comparative efficacy sources for sotatercept vs selexipag as discussed in the company response above, scenarios analyses are provided to test the uncertainty. The source applies to both the short and long-term transition probabilities. In addition, scenarios are tested reducing the RR long-term treatment effect by 50% and 60% in line with the EAG base case. However, as discussed above, the company ascertain that there is no available evidence to support a reduction in long-term RR and therefore no reduction is applied in the base case.
Selexipag long-term transitions: MAIC + 50% RR reduction	
Selexipag long-term transitions: MAIC + 60% RR reduction	
No PCA dose adjustment between selexipag-naïve and selexipag-experienced patients	The company updated base case applies differential doses of IV PCA for patients depending on whether they are selexipag-naïve or selexipag-experienced based on clinical expert feedback to the company. A scenario is provided whereby the PCA dose is equal across selexipag-naïve and selexipag-experienced patients.
IV PCA disutility scenario 1	The company base case includes an IV PCA disutility for all patients who are IV PCAs, based on the value reported in the literature. Two scenarios, as discussed in the company response above, are provided testing the impact of different assumptions. Given that the updated company model allows for improvement in risk status upon initiation of PCAs, the company maintain that IV PCA disutility should be included in the economic model base case to account for the substantial quality of life impact and administration burden these treatments have.
IV PCA disutility scenario 2	
National Audit of Pulmonary Hypertension baseline characteristics used	The base case uses baseline characteristics from the STELLAR trial as this most closely matches the population that are anticipated to receive sotatercept in clinical practice. A scenario is tested using UK baseline characteristics, age and male/female ratio, from the National Audit of Pulmonary Hypertension representing all patients with PAH regardless of PAH type or risk status. The company maintain that the STELLAR trial baseline characteristics are the most appropriate to reflect the population in this decision problem.

Table 18: Scenario analysis results for revised company base case

Scenario	Incremental costs	Incremental LYs	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Base case			**4.09**		
Mortality approach: Dependent model using Gompertz			4.88		
Selexipag long-term transitions: within-trial analysis, no RR reduction			4.22		
Selexipag long-term transitions: within-trial analysis + 50% RR reduction			3.75		
Selexipag long-term transitions: MAIC + 50% RR reduction			3.47		
Selexipag long-term transitions: MAIC + 60% RR reduction			3.35		
No PCA dose adjustment between selexipag-naïve and selexipag-experienced patients			4.09		
IV PCA disutility scenario 1			3.87		
IV PCA disutility scenario 2			4.07		
National Audit of Pulmonary Hypertension baseline characteristics used			2.55		

ICER, incremental cost-effectiveness ratio; IV, intravenous; LY, life-year; MAIC, match-adjusted indirect treatment comparison; PCA, prostacyclin analogue; QALY, quality-adjusted life-year; RR, relative risk

Table 19. Scenario analysis results for revised company base case with sotatercept drug wastage costs included

Scenario	Incremental costs	Incremental LYs	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Base case with sotatercept drug wastage costs included			**4.09**		
Mortality approach: Dependent model using Gompertz			4.88		
Selexipag long-term transitions: within-trial analysis, no RR reduction			4.22		
Selexipag long-term transitions: within-trial analysis + 50% RR reduction			3.75		

Selexipag long-term transitions: MAIC + 50% RR reduction	████	██	3.47	████	████
Selexipag long-term transitions: MAIC + 60% RR reduction	████	██	3.35	████	████
No PCA dose adjustment between selexipag-naïve and selexipag-experienced patients	████	██	4.09	████	████
IV PCA disutility scenario 1	████	██	3.87	████	████
IV PCA disutility scenario 2	████	██	4.07	████	████
National Audit of Pulmonary Hypertension baseline characteristics used	████	██	2.55	████	████

ICER, incremental cost-effectiveness ratio; IV, intravenous; LY, life-year; MAIC, match-adjusted indirect treatment comparison; PCA, prostacyclin analogue; QALY, quality-adjusted life-year; RR, relative risk

Appendix C: Guide to model adaptation for EAG

The revised company model contains a number of new settings to implement the updates discussed in the responses above.

- The capacity for patients to improve upon initiation of PCA is enable through two additional switches (one for each arm) on the Clinical inputs sheet (rows 20, 22)
- To enable the weighting of costs and QALYs in the sotatercept intermediate-high risk state (39.9% of patients also receive a PCA), a “Run model” macro has been added. This must be activated every time model inputs are changed
- A switch to adjust the PCA doses between selexipag-naïve and selexipag-experienced patients has been added to Direct cost inputs AA14
- The exclusion of sotatercept wastage costs is controlled by the “Include cost of sotatercept drug wastage” dropdown in Direct cost inputs I14
- The scenarios for IV disutility assumptions are controlled by a new dropdown on Patient HRQoL inputs E27

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Additional clarification questions following the company's Draft Guidance response

1. Within the model, the probability of improvement from the intermediate-low to low risk state on sotatercept+PD5i+ERA was previously restricted to the first 24 weeks given the lack of long-term data. This was not previously questioned by the committee. However, this 24 week threshold has now been removed in the revised base care without explanation in the response document, and the EAG are unsure if the change is intentional. Could the company please clarify and either revise or justify the change in approach and the clinical validity of the consequent model projections.

MSD Response:

The company confirm that the removal of the no improvement beyond week 24 in the sotatercept arm only, resulting in continued improved from intermediate-low to low risk for the entire model time horizon, was a result of an error in the updated formula to the 'Transition probabilities' sheet columns V and I.

The cap on the improvement in risk status beyond the first 24 weeks was applied in the original company base case as a conservative assumption to reflect the availability of data from the pivotal STELLAR trial and the lack of longer-term evidence to support an extended period of improvement. However, longer-term data available from the most recent interim analysis of the SOTERIA extension study up to 108 weeks is supportive of extending the cap in no improvement to reflect the totality of the evidence available to date. The original company submission included data from the first interim analysis of SOTERIA (Company Evidence Submission Table 56) demonstrating no evidence of a waning in the treatment effect of sotatercept. Longer-term data available from the most recent datacut of SOTERIA (15th August 2024), as presented in Table 1, demonstrates that ███ participants have reached, and have a non-missing risk status assessment, at week 108; of those, ███ participants (███) showed a risk status improvement or maintenance relative to STELLAR week 24.

Table 1. Analysis of the proportion of risk status improvement or maintenance relative to STELLAR Week 24

Endpoint	Sotatercept (N[a] = 153)		
	n/N[b]	%	(95% CI)[c]
Risk Status Improvement or Maintenance Relative to STELLAR Week 24			
Week 36	███	███	███
Week 52 (Year 1)	███	███	███
Week 108 (Year 2)	███	███	███

Database Cutoff Date: STELLAR (MK-7962-003) - 06DEC2022 / SOTERIA (MK-7962-004) - 15AUG2024
Abbreviations: CI, confidence interval; n, Number of participants with improvement or maintenance of the endpoint assessment at a given timepoint relative to week 24.
[a] Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept. 154 participants from STELLAR sotatercept arm enrolled to SOTERIA, but one participant discontinued before

receiving study dose in SOTERIA.
[b] Participants with non-missing endpoint assessment at both given timepoint and STELLAR week 24. Row percentages are based on this number.
[c] Derived based on the Clopper-Pearson method (exact binomial distribution)

The company revised base case model includes the ability for improvement in risk status from intermediate-low to low up to 108 weeks. Whilst the data from SOTERIA supports continued risk status improvement or maintenance in the sotatercept arm, the company have allowed this improvement for the selexipag arm equally to prevent any bias in favour of sotatercept. Clinical expert feedback to the company support that the main efficacy benefit of selexipag is in maintenance, rather than improvement, of risk-status and therefore this assumption is likely conservative. This extension in the ability for improvement is only applied to the intermediate-low health state to reflect the health states where sotatercept and selexipag are given on top of dual background therapy and have capacity to improve risk status. Improvement in risk status upon progression to higher risk-states, where PCA is initiated, is modelled separately as outlined in the company DG response.

The impact of extending the cap of improvement in risk status from 24 weeks to 108 weeks is explored in scenario analyses presented in Appendix A.

The implication of extending the no improvement cap on modelled outcomes is increasing the time spent in lower risk health states (low and intermediate-low), presented in Figure 1 and Figure 2. As discussed in MSD's response to the draft guidance (comment number 5; long-term transition probabilities), no appropriate data was identified from the literature reporting risk-status distribution over time in intermediate-low risk patients initiating selexipag. In addition to the challenges with comparing to the GRIPHON primary endpoint, the lack of appropriate real-world data for selexipagmakes validation of clinical outcomes incredibly difficult and largely unfeasible. a. Regardless, a comparison of the percentage of patients alive in low or intermediate-low health states in the selexipag arm at landmark timepoints is presented in Table 2.

Figure 1. Modelled alive in low or intermediate-low and OS for sotatecept and selexipag arms – no improvement from intermediate-low capped at 24 weeks

Abbreviations: OS, overall survival

Figure 2.Modelled alive in low or intermediate-low and OS for sotatecept and selexipag arms – no improvement from intermediate-low capped at 108 weeks

Abbreviations: OS, overall survival

Table 2. Modelled landmark projections for alive in low or intermediate-low risk health states in selexipag arm (company revised base case)

	% Alive in low or intermediate-low state in selexipag arm (months)						
	10	20	30	40	50	60	80
No improvement from intermediate-low to low risk capped at 108 weeks (base case)							

No improvement from intermediate-low to low risk capped at 24 weeks (scenario)	■	■	■	■	■	■	■

2. The transition probabilities applied in the model after the initiation of IV PCA in the intermediate-high risk tunnel state, following progression on selexipag, do not appear to equate with those stated in the company consultation response document (Table 1). They appear in the model to be based on Roman et al. 2012 rather than the placebo arm of STELLAR. This seems more appropriate, but could the company please clarify their indented source.

MSD Response:

The company interpret the wording of 'initiation of IV PCA' as 'initiation of PCA' as the treatment mix includes oral PCA administration in addition to IV. The revised company model includes the ability for patients to improve risk status following the initiation of PCA in either the sotatercept or selexipag arm which applies to all patients receiving PCA regardless of administration type.

The company confirm that the intended source to inform the transition probabilities in the selexipag arm following the initiation of PCAs in the intermediate-high risk tunnel state, and subsequent 'history on or off PCA' health states is the SoC arm from STELLAR as per the company consultation response document (Table 1). The company considered both sources, STELLAR SoC and Roman et al. 2012, to inform the transition probabilities for the 'history of or on PCA' health states, but concluded that STELLAR SoC would be the most appropriate with the rationale being three-fold:

1) In line with EAG and committee preferred assumptions, upon progression to higher-risk states than intermediate-low, selexipag is discontinued and PCA treatment is initiated in 85% of patients based on the expected proportion eligible and accepting; the remaining 15% of patients in the model remain on selexipag in addition to dual background therapy. As discussed in MSD's DG response (comment 7; no clinical improvement after PG12 initiation), the SoC arm from STELLAR (Week 12-24) better reflects the efficacy of a patient group receiving a combination of treatments including selexipag, inhaled and IV PCA.

2) Transitions from the SoC arm in STELLAR, as presented in the company consultation response document (Table 1), represent the probability of change in risk status in line with the risk-status based model structure. Where data is available and appropriate, risk-status based transition probabilities are preferred to best align with the risk-status based model structure.

   In contrast, transition probabilities from the Roman et al. 2012 publication are based on WHO functional class. In addition, the Roman et al. 2012 publication does not report transition probabilities from WHO FC II and therefore were Roman et al. 2012 to be used as the primary source, transition probabilities from the intermediate-low risk state history off or on PCA in the model would require an alternative source.

3) Furthermore, the use of Roman et al. 2012 to inform the transition probabilities for initiating PCAs in the intermediate high-risk tunnel state lacks face validity when comparing the applicable health state transition probabilities in the selexipag arm versus the sotatercept arm.

   To illustrate this point, the transition probabilities from the intermediate-high risk tunnel state using both the Roman et al. 2012 and STELLAR SoC (Week 12-24) values for the selexipag arm are presented in the table below for comparison to relevant transition probabilities in the sotatecept arm based on STELLAR (Week 12-24). For example, were Roman et al. 2012 to be used as the source for transition from the intermediate-high risk tunnel state to intermediate-low risk (history of or on PCA) then probability of improvement for a patient initiating PCA in addition to dual background therapy (ie. PDE5i+ERA+PCA) is greater than for a patient initiating PCA in addition dual background therapy plus sotatercept (PDE5i+ERA+sotatercept+PCA).

Table 3. Transition probabilities from the intermediate-high tunnel state

To: From:	Intermediate-low risk(history of or on PCA)	Source
Intermediate-high risk (tunnel state)	0.347	Roman et al. 2012
	████	STELLAR SoC arm (Week 12-24)
	████	STELLAR sotatercept arm (Week 12-24)

PCA, prostacyclin analogue; SOC, standard of care

In conclusion, the company maintain that the SoC arm from STELLAR is most appropriate to inform transition probabilities for the intermediate-high risk tunnel state and intermediate low, intermediate-high and high-risk ‘history of or on PCA’ health states.

3. We note that the revised model structure removes the chance of being off treatment in the intermediate-low risk state following lack of response to sotatercpet at 24 weeks. Could the company please clarify why this change has been made.

MSD Response:

In order to allow for increased complexity in the company economic model to enable risk-status improvement following the initiation of PCAs in both the sotatercept and selexipag arms, the ability for patients initiating sotatercept in intermediate-low but showing no response at 24 weeks to discontinue treatment had to be removed from the economic model. If this functionality was to remain, further additional ‘off-sotatercept’ health states would need to be added into the model to reflect

improvement in risk status for those patients initiating PCA in intermediate-high risk 'off-sotatercept'. These patients would be following a treatment pathway that is distinct from those that remain on sotatercept and therefore would need to be treated separately within the model.

The initial company model, prior to allowing for improvement on PCA, showed that only a small proportion of patients would be classified as non-responders (███), and thus discontinue sotatercept at week 24, and removing this functionality had only a small impact on the model results (see Company Evidence Submission, Table 117). Thus, the company concluded that the level of additional complexity that would be required to implement this outweighed any potential benefits given that this is not a significant driver of cost-effectiveness. Furthermore, this simplification favours the selexipag arm as the result of removing the opportunity for patients to discontinue sotatercept in non-responders at 24 weeks means that sotatercept costs will be overestimated.

The company emphasise that this was only removed to comply with the committee request to allow improvement upon PCA initiation, maintaining that this insight into non-responders remains valid and should be taken into consideration.

███**MSD Response:**

███

Of note, an error has been found in the 'Wastage' sheet of the company model relating to the calculation of sotatercept drug acquisition costs with wastage based on the patient distributions. Cells F24:G27 have been updated to reflect the correct vial sizes in line with Table 100 in the company evidence submission so that a patient in the 67.5kg – 88.9kg weight band receives 1 x 60 mg vial and a patient in the 89.0kg – 131.7kg weight band receives 2 x 45 mg vials. Previously these were inverted resulting in an overestimation of sotatercept drug acquisition costs inclusive of wastage. This change impacts scenario analysis inclusive of sotatercept wastage costs.

Appendix A: Revised Company base case results

Please note the company has updated their base case results, and these are presented in this appendix.

The base case has been revised as follows:

- Short-term transition probabilities for selexipag: use MAIC vs. GRIPHON point estimates for relative risk of risk status improvement/worsening
- Long-term transition probabilities for selexipag: application of MAIC vs. GRIPHON point estimates for relative risk of risk status improvement/worsening, with no reduction in treatment applied

- Model allows risk status improvement from intermediate-low to low risk until week 108 in both the sotatercept and selexipag arms
- Model structure updated to allow risk status improvement for patients who initiate PCA (transition probabilities in the selexipag arm informed by STELLAR SoC arm Week 12-24; transition probabilities in the sotatercept arm informed by STELLAR sotatercept arm Week 12-24)
- PCA initiation in the sotatercept arm: 39.9% of patients receive PCA when they progress to intermediate-high risk
- Proportion of PCA use: 85% of patients in intermediate-high risk and high risk in the selexipag arm, and in high risk in the sotatercept arm receive PCA. The remaining patients remain on their initial treatment.
- IV PCA dose: EAG doses for epoprostenol and treprostinil applied to patients initiating in sotatercept intermediate-high risk state, higher doses of 35ng/kg/min and 70ng/kg/min, respectively, used in selexipag intermediate-high risk state. Equivalent dosing in high risk state across arms.
- Hospitalisation-related disutility: updated to EAG value
- Hospitalisation HR between sotatercept and selexipag set equal to 1

Each of these changes has been made in an editable copy of the economic model, so that all changes can be verified against the originally submitted results, and both the company's and the EAG's original base-case analyses can be re-estimated within the same file, for transparency. Appropriate sensitivity and scenario analyses are also provided below.

Deterministic model results

The revised company base case results, for initiating sotatercept + dual background therapy vs. selexipag + dual background therapy in intermediate-low risk patients, incorporating the updated settings specified in the responses above are presented in Table 4.

Sotatercept in combination with dual therapy accumulated ████ of costs and ██ QALYs and selexipag in combination with dual therapy accumulated ████ of costs and ██ QALYs. The resulting ICER is ████ per QALY, inclusive of the 1.2 QALY severity modifier for which sotatercept qualifies.

Table 4. Deterministic base case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Sotatercept + dual therapy	███	███	███	███	███	2.70	███	███

Selexipag + dual therapy	████	████	████	-	-	-	-	-

ICER: incremental cost-effectiveness ratio; LYG: life-years gained; QALY: quality-adjusted life-years

Table 5. Deterministic net health benefit and net monetary benefit

Technologies	Incremental costs (£)	Incremental QALYs	NHB at £20,000	NHB at £30,000	NMB at £20,000	NMB at £30,000
Selexipag + dual therapy	████	2.70* 1.2 = 3.23	████	████	████	████

NHB: net health benefit; NMB: net monetary benefit; LYG: life-years gained; QALY: quality-adjusted life-years

Exploring uncertainty

Deterministic sensitivity analysis

Figure 3. Tornado diagram

BGT, background therapy; HCRU, healthcare resource use; IH, intermediate-high; IL, intermediate-low; IV, intravenous; PAH, pulmonary arterial hypertension; QALY, quality-adjusted life-years; RR, relative risk; RS, risk status

Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was run for 1,000 iterations. A weighted probabilistic ICER is generated using the proportion of iterations that met the 1.2 severity modifier threshold.

The parameters varied in the PSA and the nature of the uncertainty around their value are described in Table 13 of the company response to DG consultation.

The mean summary results from the probabilistic analysis are presented in Table 6 below. Graphical results including the cost-effectiveness plane (incremental costs vs. incremental QALYs) and the cost-effectiveness acceptability curve (CEAC) are shown below in Figure and Figure respectively. Sotatercept was demonstrated a probability of being a cost-effective treatment initiated in intermediate-low risk patients compared to selexipag in ████ and ████ of iterations at a willingness-to-pay threshold of £20,000/QALY and £30,000/QALY, respectively.

Table 6. Probabilistic base case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)	ICER weighted according to severity modifier (£/QALY)
Severity modifier							███ of iterations	███ of iterations	Weighted ICER
Sotatercept + dual therapy	███	███	███	███	███	2.64	███	███	███
Selexipag + dual therapy	███	███	███	-	-	-	-		-

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

Figure 4. Cost-effectiveness plane scatter plot (1,000 iterations) – Incremental Costs and QALYs (severity modifier applied)

QALY, quality-adjusted life-year; WTP, willingness-to-pay

Figure 5. Cost-effectiveness acceptability curve (1,000 iterations) (severity modifier applied)

BGT, background therapy; QALY, quality-adjusted life-year

Scenario analysis

Table 7. List of scenario analysis

Scenario	Rationale
Mortality approach: Dependent model using Gompertz	The dependent model, jointly fitting a parametric model to each risk status Kaplan-Meier curve from Rosenkranz, to estimate long-term overall survival is tested in scenarios. The evolution of risks over time, as provided by the Gompertz distribution, was judged to be the most appropriate alternatives to the base case choice and therefore included in the scenario analyses.
Selexipag long-term transitions: within-trial analysis, no RR reduction	Acknowledging limitations associated with both the MAIC and WTA comparative efficacy sources for sotatercept vs selexipag as discussed in the company response above, scenarios analyses are provided to test the uncertainty. The source applies to both the short and long-term transition probabilities. In addition, scenarios are tested reducing the RR long-term treatment effect by 50% and 60% in line with the EAG base case. However, as discussed above, the company ascertain that there is no available evidence to support a reduction in long-term RR and therefore no reduction is applied in the base case.
Selexipag long-term transitions: MAIC + 50% RR reduction	
Selexipag long-term transitions: MAIC + 60% RR reduction	
Improvement in risk status from intermediate-low risk capped at 24 weeks for sotatercept and selexipag	The updated company base case applies a rule of no improvement from the intermediate-low risk health states beyond week 108, supported the latest available data from the SOTERIA trial demonstrating no waning in treatment effect. This is applied to both arms. A scenario analysis explores this cap remaining at 24 weeks.
Improvement in risk status from intermediate-low risk capped at 24 weeks for selexipag arm and 108 weeks for sotatercept arm	The company updated base case applies a no improvement rule beyond week 108 to both the sotatercept and selexipag arms in the economic model. Extending the possibility of improvement in both arms is a conservative assumption as evidence from SOTERIA is available only to support no treatment waning effect in

Scenario	Rationale
	participants receiving sotatercept. Therefore a scenario explores the no improvement in risk status cap in the selexipag arm remaining at 24 weeks.
No PCA dose adjustment between selexipag-naïve and selexipag-experienced patients	The company updated base case applies differential doses of IV PCA for patients depending on whether they are selexipag-naïve or selexipag-experienced based on clinical expert feedback to the company. A scenario is provided whereby the PCA dose is equal across selexipag-naïve and selexipag-experienced patients.
IV PCA disutility scenario 1	The company base case includes an IV PCA disutility for all patients who are IV PCAs, based on the value reported in the literature. Two scenarios, as discussed in the company response above, are provided testing the impact of different assumptions. Given that the updated company model allows for improvement in risk status upon initiation of PCAs, the company maintain that IV PCA disutility should be included in the economic model base case to account for the substantial quality of life impact and administration burden these treatments have.
IV PCA disutility scenario 2	
National Audit of Pulmonary Hypertension baseline characteristics used	The base case uses baseline characteristics from the STELLAR trial as this most closely matches the population that are anticipated to receive sotatercept in clinical practice. A scenario is tested using UK baseline characteristics, age and male/female ratio, from the National Audit of Pulmonary Hypertension representing all patients with PAH regardless of PAH type or risk status. The company maintain that the STELLAR trial baseline characteristics are the most appropriate to reflect the population in this decision problem.

Table 8. Scenario analysis results for revised company base case including sotatercept wastage

Scenario	Incremental costs	Incremental LYs	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Base case	████	████	**2.70**	████	████
Mortality approach: Dependent model using Gompertz	████	████	3.21	████	████
Selexipag long-term transitions: within-trial analysis, no RR reduction	████	████	2.88	████	████
Selexipag long-term transitions: within-trial analysis + 50% RR reduction	████	████	2.07	████	████
Selexipag long-term transitions: MAIC + 50% RR reduction	████	████	4.08	████	████

Selexipag long-term transitions: MAIC + 60% RR reduction			4.26		
No improvement from IL risk capped at 24 weeks			2.16		
No improvement from IL risk capped at 24 weeks for selexipag and 108 weeks for sotatercept			2.79		
No PCA dose adjustment between selexipag-naïve and selexipag-experienced patients			2.70		
IV PCA disutility scenario 1			2.51		
IV PCA disutility scenario 2			2.70		
National Audit of Pulmonary Hypertension baseline characteristics used			1.88		
No sotatercept drug wastage costs			2.70		

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

	Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly. The Appraisal Committee is interested in receiving comments on the following: • has all of the relevant evidence been taken into account? • are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? • are the provisional recommendations sound and a suitable basis for guidance to the NHS? NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the preliminary recommendations may need changing in order to meet these aims. In particular, please tell us if the preliminary recommendations: • could have a different impact on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology; • could have any adverse impact on people with a particular disability or disabilities. Please provide any relevant information or data you have regarding such impacts and how they could be avoided or reduced**.**
Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):	Dr Daniel Knight, Royal Free London NHS Foundation Trust

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months. [Relevant companies are listed in the appraisal stakeholder list.] Please state: • the name of the company • the amount • the purpose of funding including whether it related to a product mentioned in the stakeholder list • whether it is ongoing or has ceased.	I have received consultancy meeting fees from MSD for meetings about sotatercept on 12/11/24 (for £672) and 06/02/25 (for £224). In the last 12 months, I have also received speaker fees from Janssen Pharmaceuticals (comparator treatment, selexipag) on 13/11/24 (for £447.93) and 10/12/24 (for £1045.47). All of these disclosures were for individual meetings or talks (i.e. not ongoing).
Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	I have no disclosures with respect to the tobacco industry.
Name of commentator person completing form:	Dr Daniel Knight

Comment number	**Comments** Insert each comment in a new row. Do not paste other tables into this table, because your comments could get lost – type directly into this table.
Example 1	We are concerned that this recommendation may imply that …………..
1	**Page 3** – "*The company asked for sotatercept to be considered only for people who have an intermediate-low-risk status to reflect the populations in the clinical trials. This does not include everyone it is licensed for*"

NICE National Institute for Health and Care Excellence

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments the end of 17 September 2025. Please submit via NICE Docs.

	Page 7 – "*The committee acknowledged that this treatment was positioned by the company only for PAH with an intermediate-low-risk status*" **Page 8** – "*The company explained that their positioning of sotatercept for ESC/ERS intermediate–low-risk PAH is based on the current marketing authorisation and available study data*" I agree with NICE that sotatercept should also be considered for patients with PAH and intermediate-high-risk and high-risk statuses. Indeed, we have a significant unmet treatment need in this high morbidity and mortality patient population. However, my impression from the meeting was that it was the lack of a feasible treatment comparator in these groups that limited the scope of the case submission to intermediate-low risk patients. The data from Zenith (https://pubmed.ncbi.nlm.nih.gov/40167274/), which was not in the original submission, is compelling for the use of sotatercept in high-risk patients with PAH receiving the maximum tolerated dose of background PAH therapy. This was placebo-controlled (rather than comparing with another PAH therapy). However, it would be difficult to compare any drug with intravenous epoprostenol, as this is usually the 'exit strategy' for most PAH clinical trials.
2	**Page 7** – "*Some people need to take their PCA treatment with them everywhere, which can risk infection*" Intravenous PCA is a continuous 24/7 infusion through an indwelling venous catheter. It cannot be stopped as it has a half-life of approximately 4-5 minutes. Therefore, the continuous nature of this treatment (and consequent infection risk of an indwelling venous catheter) applies to all patients, all of the time.
3	**Page 8** – "*The clinical expert explained that 38% of people with PAH are intermediate–high risk and 17% are high risk*" To specify, these proportions are at first follow-up after 3-4 months of initial treatment based on COMPERA data.
4	**Page 14** – "*In the model, people move to the intermediate–high and high-risk states on disease progression. On entering these states, intravenous PCA is started as a subsequent treatment*" As mentioned further on page 16 of the draft guidance, the problem with this assumption is that intravenous PCA is not necessarily suitable for all patients. For example, they may be unable to be safely trained to reconstitute, administer and manage the treatment (for a variety of reasons). So, we have a proportion of patients at the highest mortality risk who are unable to access this treatment for other reasons, hence the therapy gap we have in PAH. Whilst there has been a paradigm shift in our approach to PAH therapy (i.e. generally earlier introduction of PAH therapy to attain lowest risk status at the earliest opportunity), a historical review of a subset of patients from the REVEAL registry (https://pubmed.ncbi.nlm.nih.gov/24035189/) showed that: (i) among patients who had a PAH-related death, only 56% were receiving intravenous prostacyclin before death; (ii) in WHO FC 4 patients who had a PAH-related death, about one-third were not on parenteral prostanoid. This may be, at least in part, due to the difficulties of intravenous treatment that I mentioned above.
5	**Page 15** – "*The clinical expert agreed that clinical improvement can be seen after starting intravenous PCAs. But, they stated that the main benefit of intravenous PCAs is stability in the intermediate–high or high-risk states*" I apologise if I confused two issues but I would like to clarify this comment as I think it might have confused two separate PAH treatments: - Intravenous PCA has an evidence-based benefit of reduced mortality in PAH along with concomitant clinical improvement. - Selexipag in the intermediate-low risk PAH group does not have an evidence-base for symptom or exercise tolerance improvement. Instead, selexipag has more of a benefit in terms disease stability

	in the intermediate-low risk state (as the majority of primary endpoint events in GRIPHON were non-fatal clinical worsening events).
6	**Page 22** – "*It noted that while the company positioned it for people with intermediate–low-risk status, there is evidence published since the company submission that suggests this could benefit people in intermediate–high- or high-risk states. The committee expressed that this option should be explored before the next committee meeting*" I am very much in favour of exploring the scope for sotatercept in England for patients with PAH in the intermediate-high and high-risk states, as well as intermediate-low risk patients. There are compelling data from studies published since the original submission: (i) HYPERION (https://pubmed.ncbi.nlm.nih.gov/41025556/) - phase 3 trial of patients with WHO FC II or III PAH who had received the diagnosis less than 1 year earlier, had an intermediate or high risk of death, and were receiving double or triple background therapy. There was a lower risk of clinical worsening with sotatercept versus placebo. (ii) ZENITH (https://pubmed.ncbi.nlm.nih.gov/40167274/) - phase 3 trial of patients with WHO FC 3 or 4 PAH and a high 1-year risk of who were receiving the maximum tolerated dose of background therapy to receive add-on sotatercept or placebo. Sotatercept resulted in a lower risk of a composite of death from any cause, lung transplantation or hospitalization for worsening PAH than placebo. I also have a manuscript currently under review demonstrating significant benefit in 6-minute walk, PAH risk score and cardiac MRI metrics of right heart size and function in 20 patients who were intermediate-high or high risk by ESC/ERS 2022 despite triple PAH therapies including IV PCA (I am happy to share these data with the committee, although I appreciate it currently remains unpublished at the time of writing). Overall, there is a significant unmet therapeutic need in this high-morbidity, high-mortality disease area, and I believe sotatercept offers a meaningful opportunity to address it. Many patients, especially those at intermediate-high and high mortality risk, currently lack effective long-term treatment options and so the burden of disease remains unacceptably high. We are falling behind the USA and several European countries in adopting this innovative therapy. Timely access to sotatercept in the UK could help close this gap and improve outcomes for patients who urgently need better treatment solutions.

Insert extra rows as needed

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about funding from the company and links with, or funding from, the tobacco industry.
- Combine all comments from your organisation into one response. We cannot accept more than one set of comments from each organisation.
- Do not paste other tables into this table – type directly into the table.
- In line with the NICE Health Technology Evaluation Manual (sections 5.4.4 to 5.4.21), if a comment contains confidential information, it is the responsibility of the responder to provide two versions, one complete and one with the confidential information removed (to be published on NICE's website), together with a checklist of the confidential information. Please underline all confidential information, and separately highlight information that is submitted as 'confidential [CON]' in turquoise, and all information submitted as 'depersonalised data [DPD]' in pink. If confidential information is submitted, please submit a second version of your

comments form with that information replaced with asterixis and highlighted in black.

- Do not include medical information about yourself or another person from which you or the person could be identified.
- Do not use abbreviations.
- Do not include attachments such as research articles, letters or leaflets. For copyright reasons, we will have to return comments forms that have attachments without reading them. You can resubmit your comments form without attachments, it must send it by the deadline.
- If you have received agreement from NICE to submit additional evidence with your comments on the draft guidance document, please submit these separately.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Single Technology Appraisal

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Comments on the draft guidance received through the NICE website

Name	██████████
Organisation	Janssen
Conflict	Janssen are a comparator company
Comments on the DG:	

Committee Discussion - Section 3.3 - Current treatment of PAH

In accordance with the DG, we support the appraisal of sotatercept in a way that best addresses clinical burden and unmet need of UK patients — including, those with WHO FC III-IV disease, who are often in the intermediate-high or high-risk categories and may require quadruple-therapy in order to achieve optimal clinical outcomes.

Established PAH therapies target three key biological pathways; nitric oxide (PDE-5 inhibitors/sGC stimulators), endothelin (endothelin receptor antagonists), and prostacyclin (parenteral prostanoids such as IV epoprostenol, or oral IP receptor agonists such as selexipag). More recently, modulation of the activin/BMP pathway such as with sotatercept has emerged as a 4th additional therapeutic pathway, reflected in the WSPH treatment recommendations (https://publications.ersnet.org/content/erj/early/2024/06/13/1399300301095-2024).

Clinical trial data, including PULSAR, STELLAR, HYPERION, and ZENITH, demonstrate that a substantial proportion of sotatercept-treated patients are relatively more severe and higher risk PAH patients on a background of triple PAH therapies, indicating that sotatercept could play a role as an add-on therapy to triple PAH therapy.

Committee Discussion - Section 3.5 - Positioning of sotatercept

We believe that aligning the assessment of sotatercept with the most comprehensive and robust clinical evidence is essential to support well-informed decision-making. The current positioning may not be fully aligned with the available data or generalisable to the UK population, and therefore we support the view that this may benefit from further reflection.

Sotatercept's evidence supports patient benefit in patients with more advanced disease stages—namely, patients with WHO III-IV functional class, who are likely to be classified within the intermediate-high or high risk 2022 ESC/ERS 4-strata risk categories as per the proposed WSPH treatment algorithm recommendations. The clinical trials—PULSAR, STELLAR,

HYPERION, and ZENITH—include a significant proportion of sotatercept-treated patients on background triple therapy with a relatively long time period since diagnosis (~9 years in STELLAR). This may indicate a role of sotatercept as an add-on to ensure comprehensive pathway coverage. Specifically:

- PULSAR: n=106; WHO FC III: n=50 (47%); background triple-therapy: n=59 (56%)
- STELLAR: n=323; WHO FC III: n=84 (52%); background triple-therapy: n=198 (61%)
- HYPERION: n=320; WHO FC III: n=252 (78.8%); Background triple-therapy: n=89 (27.8%)
- ZENITH: Total: n=172; WHO FC III: n=128 (74%), WHO FC IV: n=44 (26%); background Triple-therapy: n=124 (72%)

Consideration of the full evidence package may be aligned with clinical desire to use this medicine as part of quadruplet therapy in a more intuitive addition to the current treatment pathway. We understand and support the view expressed in the DG that this would enable the Committee to make a fully informed, sound, and suitable recommendation.

To this end, we believe that consideration of these key data points and acknowledging the severity of disease in later stages would help ensure that the evidence base truly reflects the potential benefits of sotatercept for those who need it. We hope that a comprehensive review of these data can support a decision that aligns with the goal of offering the most effective options to patients with significant unmet needs.

Committee Discussion - Section 3.7 - Indirect treatment comparisons

Over the past several decades, the demographic profile of PAH patients in the UK and Europe has changed markedly. Contemporary PAH patient registry data (e.g. COMPERA, French registry, ASPIRE etc.) show the average European/UK PAH patient is commonly older (~60 years old) and often has comorbidities (e.g. cardiovascular disease) making clinical management more complex. This real-world profile underlines the importance of evidence in broad, diverse patient populations that closer reflects NHS practice. (https://publications.ersnet.org/content/erj/51/5/1800629, https://www.thelancet.com/journals/lanres/article/PIIS2213-2600(22)00097-2/fulltext & https://publications.ersnet.org/content/erj/39/4/945)

Selexipag has a large evidence base as a relatively mature PAH therapy, composing of a breadth of clinical trial (GRIPHON) and real-world data, and includes relevant sub-group analyses of GRIPHON that align with current patient profiles. The largest trial in PAH patients to date is GRIPHON — a global, event-driven, double-blind, randomised, placebo-controlled Phase III study (n=1,156) – and includes a composite morbidity/mortality (M/M) primary endpoint of over ~3 years of selexipag treatment. In GRIPHON, selexipag reduced the risk of morbidity/mortality events by 40% versus placebo (HR

0.60; 99% CI 0.46-0.78; p<0.001). Moreover, the GRIPHON trial included diverse subgroups of PAH patients typically treated in the UK providing evidence of efficacy of selexipag in various published post-hoc analyses.

The 10-year follow-up open-label extension data of GRIPHON was recently published reporting PAH patients treated with selexipag for 10 years had a ~60% survival rate and corresponding to a total selexipag exposure of 2105.5 patient-years, exceeding the UK the National Audit data median survival of ~6 years of overall PAH. From the real-world perspective, evidence of outcomes of selexipag in real-world clinical practice is demonstrated via the European selexipag post-approval safety study EXPOSURE (n=~700) and from the US SPHERE (n=~800). Since 2015 when GRIPHON was first published, up to November 30, 2023, an estimated 46,452 PAH patients have been exposed to commercial selexipag worldwide, and in the UK the National Audit data show over 200 PAH patients with a selexipag prescription as far back as 31st March 2020, increasing steadily to 460 in the most recent 2024 data. (https://pmc.ncbi.nlm.nih.gov/articles/PMC11284239/#pul212403-bib-0013, https://pubmed.ncbi.nlm.nih.gov/39083197/, & https://www.jhltonline.org/article/S1053-2498(23)02027-2/fulltext S). Such long term data may be useful in informing the economic model considered in this appraisal.

As noted in the DG, the potential differences in the study population may lead to heterogeneity in the indirect comparison. Initial feasibility assessments conducted by J&J IM indicated that severe limitations in study population comparability limited the value of such methods.

In particular, the GRIPHON and STELLAR trials differed in terms of the included patient populations with respect to age, time since diagnosis, PAH aetiology classification, WHO Functional Class, and background PAH therapy (proportions receiving none/mono/double/triple therapy). All of these characteristics could be prognostic of patient outcomes.

Given the small degree of overlap between the study populations, any adjustment in an indirect comparison is likely to be extreme and highly uncertain. However, such adjustment would be necessary to avoid incorrect estimation of relative clinical benefit and evaluate how patients can be optimally treated to ensure they experience the best outcomes possible.

Committee Discussion - Section 3.22 – Conclusion

We are aligned with the committee view that the positioning of the medicine does not reflect the full range of patient population with unmet clinical need. Use of sotatercept as part of quadruple-therapy could be beneficial, and the analyses supporting this request would be informative.

As per other comments, the additional analyses requested by NICE may well be more relevant in addressing the decision problem appropriately and in line with robust standards

Name	[redacted]
Organisation	N/A
Conflict	N/A

Comments on the DG:

Has all of the relevant evidence been taken into account?

Yes there is detailed evidence that has been submitted in support of this drug. And I understand part of the economic modelling. However as someone who has used this drug in int-low and high risk patients it is clear to me that this is a very effective drug and can rapidly change the outcome for patients with this life shortening disease. It is a shame to deny access of this treatment to patients.

Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?

Yes - but the positioning is difficult. The committee acknowledge that there is new data now out that suggests ongoing benefit in wider population of at risk patients, Hopefully a resubmission will support this.

Are the recommendations sound and a suitable basis for guidance to the NHS?

Not really. The guidance is clear cut but I think misses the point in care of an orphan disease. These patients have been underserved with therapies that vasodilator the pulmonary vasculature rather than actually trying to reverse remodel the vessel. This treatment looks to be trying to help that process and therefore the wider view should be taken that in nationally specialist centres physicians can identify those patients that would benefit the most from this therapy.

Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of race, gender, disability, religion or belief, sexual orientation, age, gender reassignment, pregnancy and maternity?

No. Just that this is an orphan disease. Therefore, the population in the UK is small and underserved.

Name	████████
Organisation	N/A
Conflict	N/A
Comments on the DG:	

Has all of the relevant evidence been taken into account?

I appreciate that this appraisal process has been started several months ago based mainly on the data from the STELLAR trial and therefore most recent data from RCT are not included. I would like to point to recently publicly available data from Phase 3 Zenith (ZENITH ClinicalTrials.gov number, NCT04896008) and HYPERION studies. These trials were performed in different patient populations.

In ZENITH (patients in FC III/IV on maximal background therapy patients treatment with sotatercept resulted in a lower risk of a composite of death from any cause, lung transplantation, or hospitalization (≥24 hours) for worsening pulmonary arterial hypertension than placebo. Results show 76% reduction in risk of major morbidity or mortality events.
https://www.nejm.org/doi/full/10.1056/NEJMoa2415160 .
In the HYPERION global, randomized, double-blind, placebo-controlled, Phase 3 clinical trial evaluating sotatercept in combination with background therapy for newly diagnosed intermediate- and high-risk pulmonary arterial hypertension patients. In January 2025 that the trial was stopped early for a final analysis, as it was deemed unethical to continue due to positive data from the overall program. The trial was a continuation of earlier research that showed sotatercept could significantly improve Pulmonary Vascular Resistance. https://www.merck.com/news/merck-announces-phase-3-hyperion-study-of-winrevair-sotatercept-csrk-met-primary-endpoint-in-recently-diagnosed-adults-with-pulmonary-arterial-hypertension-pah/

The results of the trial will be presented at ERS congress as the a late breaking presentation on 30th of September 2025. ZENITH and HYPERION studies provide crucial data for the approval and use of sotatocept for treatment for PAH in my opinion. I have been involved in management of patients with PAH and drug development since late 1990's as the researcher, consultant, opinion leader working with NICE (on the appraisal of IV epoprostenol and with Commissioning Policy for PAH drugs). I am co-author of International Guidelines (2022 ESC/ERS Guidelines on PH) and actively contribute to the works of the consecutive World Symposia on PH since 2003 including last 7WSPH in 2024. With all my experience in drug development in PAH, Sotatercept shows most effectiveness in controlling/reverting disease process out of all developed class of drugs for PAH so far. The trials were stopped early for a final analysis, as it was deemed unethical to continue due to positive data from the overall program.

This new class of drugs, a first-in-class activin signaling inhibitor, targets the pathological mechanisms underlying vascular remodelling in pulmonary arterial hypertension (PAH). HYPERION results are demonstrating that early addition of sotatercept to background therapy in a newly diagnosed, incident

PAH population delays time to clinical worsening compared to standard care alone.

This data is showing that we can prevent disease progression with treatment of sotatercept and keep patients in improved risk strata.

Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?

The summary, of clinical and cost effectiveness is interpreted based on the presented evidence and according to NICE criteria. I would like to point out that science has fast moved forward, since original documents were filed and there is new available important data and perhaps this appraisal should be updated as the effectiveness of the drug is making significant difference in patients outcome measured as trials end points including PRO.

Are the recommendations sound and a suitable basis for guidance to the NHS?

I understand health economic appraisal process which probably is easier to perform for the drugs in common condition. It is more difficult and more compels for rare/orphan diseases as PAH, and the balance between benefits to the patients, society, NHS and costs is more difficult to capture and balance apropriatly. I hope that new data and reassuring safety profile can change the balance towards recommendation for specific populations with PAH.

Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of race, gender, disability, religion or belief, sexual orientation, age, gender reassignment, pregnancy and maternity?

This area is well covered.

Sotatercept for treating pulmonary arterial hypertension [ID6163]

EAG critique of the company response to the Appraisal Consultation Document

Produced by Aberdeen HTA Group

Correspondence to Graham Scotland
Health Economics Research Unit, University of Aberdeen
2nd Floor, Health Sciences Building, Foresterhill
Aberdeen, AB25 2ZD
g.scotland@abdn.ac.uk

Date completed **13 Nov 2025**

Contains

Copyright belongs to the Aberdeen HTA Group, University of Aberdeen unless otherwise stated.

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number **NIHR166802.**

[illegible]

Following the first committee meeting for the appraisal, the Committee were minded not to recommend sotatercept for the treat pulmonary arterial hypertension (PAH) in adults with WHO Functional Class (FC) 2 to 3 to improve exercise capacity.

Rationale for the preliminary decision included:

- uncertainty regarding the long-term treatment effect of sotarercept versus selexipag, with the model applying short term relative risks to generate transition probabilities for selexipag over the modelled lifetime horizon.
- Inconsistent assumptions about treatment initiation with prostacyclin analogues (PCA) after progression.
- Lack of modelling of clinical improvement after initiation of PCA upon progression.
- Potential overestimation of IV PCA costs due to high dosing assumptions

The committee requested the following additional analyses in the draft guidance:

1. Comparative Analysis:
    - Comparison of sotatercept with PCA in intermediate–high and high-risk populations using published data (e.g., ZENITH trial).
2. Model Structure Updates to:
    - Allow for clinical improvement after starting prostacyclin analogues.
3. Statistical Adjustments:
    - Use of propensity score matching to adjust for baseline differences in the STELLAR post-hoc analysis (per NICE Technical Support Document 17).
4. Validation of Relative Risk Reductions:
    - Use alternative data sets to validate the relative risk reductions derived from the indirect treatment comparison (ITC) with GRIPHON and TRACE.
5. Updated Scenario Analyses, including:
    - Aligning population characteristics with the UK National Audit of Pulmonary Hypertension (NAPH).
    - Starting PCA analogues with sotatercept
    - different modelling approaches for overall survival in sensitivity analyses.

In response to the committee requests, the company have revised their model structure to allow for improvement in risk status following initiation of PCA and provided a new matched adjusted indirect comparison between sotatercept and selexipag based on data from STELLAR and GRIPHON.

The company have not provided a comparison between sotatercept and prostacyclin analogues in the intermediate-high and high risk groups as requested in point 1. They note that intermediate-low risk is proposed as the appropriate population in which to initiate treatment with sotatercept, not the population to whom use should be restricted. Those who improve to low risk or progress to intermediate-high risk are expected to remain on sotatercept treatment. They do not, however, provide a case for initiation of sotatercept versus initiation of infused PCA in the intermediate high or high risk states. The ZENITH trial, being placebo controlled, is not well suited to informing this comparison; with those on PCA in the intermediate-high and high risk groups being on them at baseline.

They further note the limitations of the ZENITH trial for informing the comparative efficacy of sotatercept + PDE5i+ERA+PCA versus PDE5i+ERA+PCA in the high risks state, given low numbers available to inform transitions in this post-hoc comparison (see company response, point 6). Given the low numbers stated, the EAG accept that there are too few observations to adequately inform the relative benefit of this maximal therapy (sotatercept+PDE5i+ERA+PCA) versus PDE5i+ERA+PCA. A further limitation is that sotatercept was added to a background of PDE5i+ERA+PCA in ZENITH. What is required, is the effect of initiating PCA following progression on sotatercept+PDE5i+ERA versus initiating or remaining on PCA on top of PDE5i+ERA following discontinuation of selexipag. Unfortunately, data to inform this are not available.

The company have not provided an adjusted comparison between the sotatercept plus dual background therapy and selexipag plus dual background therapy subgroups of STELLAR (as requested in point 3 above). They note that the sample size would be too small to allow for appropriate adjustments in this non-randomised comparison (see point 3 of the company response document). The EAG agree with this. Further, any adjustment would not overcome the main issue of selexipag initiation being unobserved in STELLAR, making it unsuitable for directly informing the comparative efficacy of selexipag versus sotatercept.

Ä¾☺ f9 ¾╤☺¾τ ±Exx • τ× • f¢c á Ä

In their original submission, the company presented a BUCHER ITC, comparing sotetercept with selexipag on a range of outcomes using data from the STELLAR trial and pooled data from the GRIPHON and TRACE selexipag trials. It was stated in the ACD that the committee preferred using the risk ratios for WHO functional class deterioration and improvement from this ITC for estimating short term transition probabilities for ESC/ERS risk status

deterioration and improvement on selexipag+PDE5i+ERA. The committee, however, also suggested that a matching-adjusted indirect comparison (MAIC) could have been used in this ITC to adjust for population differences between the trials.

In their response, the company have provided a new anchored MAIC, using data from the STELLAR trial adjusted to match the distribution of baseline characteristic of the cohort recruited to the GRIPHON trial (see point 3 and Appendix A of the company response document). A single trial had to be chosen for matching purposes due to a lack of reported baseline data for the pooled population of GRIPHON and TRACE.[1,2] GRIPHON was selected as the largest of the two trials (N=1,156),[1] with a more comparable population to the STELLAR trial in terms of baseline characteristics. To align the populations better, only those on monotherapy or dual background therapy at baseline in STELLAR were used; all patients in the GRIPHON trial were receiving background monotherapy or dual therapy at baseline. This step removed 61% (=198/323) of STELLAR trial patients from the analysis, leaving 125 for inclusion in the MAIC (n=65 in the sotatercept and n=60 in the placebo arm). This group was adjusted to align with the distribution of baseline characteristics reported for the GRIPHON trial population (Effective Sample Size = ███ after matching). The company note that, to maintain a reasonable effective sample size, the adjustments were restricted to those with the highest likelihood of being important treatment effect modifiers. Thus, baseline WHO diagnostic group, WHO FC, age, and six-minute walking distance were adjusted for. Following the adjustments, the company generated new risks ratios for WHO FC improvement and deterioration as presented in their response document. The notable impact is a greater relative reduction in the risk of deterioration in WHO FC at 24 weeks (███; 95%CI: ██████), compared to the previous estimate of █████████ from the unadjusted Bucher ITC. The company use their new MAIC-based risk ratios of improvement and deterioration to generate transition probabilities for selexipag relative to sotatercept in their new cost-effectiveness base case.

EAG comment on the MAIC

It is of some concern that the new MAIC only matched on selected characteristics from a larger set of potential effect modifiers. For anchored MAICs to produce unbiased estimates, all relevant treatment effect modifiers need to be measured and correctly adjusted to match the comparator trial population.[3] With a small sample, and limited availability of data on potential effect modifiers, important differences may remain between STELLAR and GRIPHON which have potential to bias the ITC.

In addition, limiting the MAIC to include the subpopulation of STELLAR on mono and dual background therapy at baseline, removes a large proportion of patients (61%) from the sample, and increases the chance of creating imbalance in prognostic factors and confounding the comparison between sotatercept and placebo in the remaining subgroup. Potential therefore remains for bias the the subsequent ITC via Bucher method. It may be further noted that subgroup analysis of the STELLAR trial (on 6MWD or NT-proBNP - see 2.8.1 of the Company Submission) doesn't appear to support background triple combination therapy as an effect modifier of sotatercept. Therefore, it is questionable that the MAIC has improved the validity of the ITC with selexipag.

The EAG are of the opinion that high uncertainty remains regarding the magnitude of benefit for sotatercept versus selexipag, particularly over the longer-term.

E♠╤╖Ϊ ╓─Ж ╤ň ҭ ╤̰ ╡̰ ҭ ě¾╖;╤¾╖ ╡¢╛

To extrapolate long-term efficacy of selexipag versus sotatercept in the model, the committee previously concluded, based on consultation with an expert at the meeting, "it would be reasonable to use the EAG mid scenario percentage of relative risk reduction to the treatment effects". This assumed 50% of the relative risk reduction estimated for sotatercept versus selexipag from the ITC with GRIPHON and TRACE, based on data at 24-26 weeks. The committee also acknowledged the need for alternative data sets to inform and validate the relative risk reduction. No further data has been identified by the company to validate or refute the longer-term efficacy of sotatercept versus selexipag. However, the company now prefer the updated relative risks obtained from the new MAIC between STELLAR and GRIPHON and apply these full effects in the long-term in their updated base case. The company suggest the application of reduced relative risks is "arbitrary and implicitly reflective of a wane in sotatercept relative efficacy which has not been observed in the open-label long-term follow-up of the SOTERIA trial".

EAG.comment.on.the.long_term.transition.probabilities

The EAG would note that as an uncontrolled open label extension study, the SOTERIA study did not assess relative efficacy of sotatercept. Whist SOTERIA may support maintenance of disease control for sotatercept, it doesn't inform the relative efficacy of sotatercept versus placebo or selexipag. The EAG scenarios reduced the relative risks of improvement/deterioration with selexipag after noting their continued long-term application in the company model resulted in pessimistic projections of disease progression and initiation of PCA in the selexipag arm that were inconsistent with the expectations of clinical experts;

i.e. they resulted ██ of patients remaining controlled on selexipag by three years in the model, when follow-up to three years in GRIPHON put this figure at approximately 60%.[1] For this reason, the EAG is concerned that the derived 24 week relative risks, are inappropriate for application to long term transition probabilities.

It may be further noted that restricting the relative effects of selexipag to be based on the ITC of WHO FC at 24-26 weeks, assumes that it is no better than placebo in preventing risks status deterioration. This is because selexipag demonstrated no benefit over placebo on WHO FC deterioration at 26 weeks in GRIPHON[1] (risk ratio= ████████████) (Company submission). Nevertheless, there is good evidence from the GRIPHON trial to support a significant reduction in the primary composite end point of death (from any cause) or a complication related to pulmonary arterial hypertension over three years of follow-up. The company model does not support this. Figure 1 below shows the projected proportion of patients remaining alive and free from progression (to intermediate-high risk) on selexipag, against the observed estimates of death or clinical worsening from the GRIPHON trial.

The EAG remain concerned that the company model substantially underestimates the clinical benefit of selexipag. In the absence of long-term data to inform the relative efficacy of selexipag versus sotatercept, it is important that model projections are scrutinised and validated against clinical expert opinion. The clinical expert consulted at the meeting supported a more conservative extrapolation, as did the clinical experts consulted by the EAG. To address these concerns, the EAG suggest an alternative approach, whereby the long-term transition probabilities applied to selexipag are calibrated so that the inferred weighted hazard ratio between treatment arms, for the progression to intermediate-high risk or death, is aligned with the estimated HR for non-fatal clinical progression or death (███) obtained from the company's original ITC between selexipag and sotatercept. This creates model projections as indicated in Figure 2 below.

Figure 1: Model projections of proportion alive and free of progression to higher risk states (company revised base case)

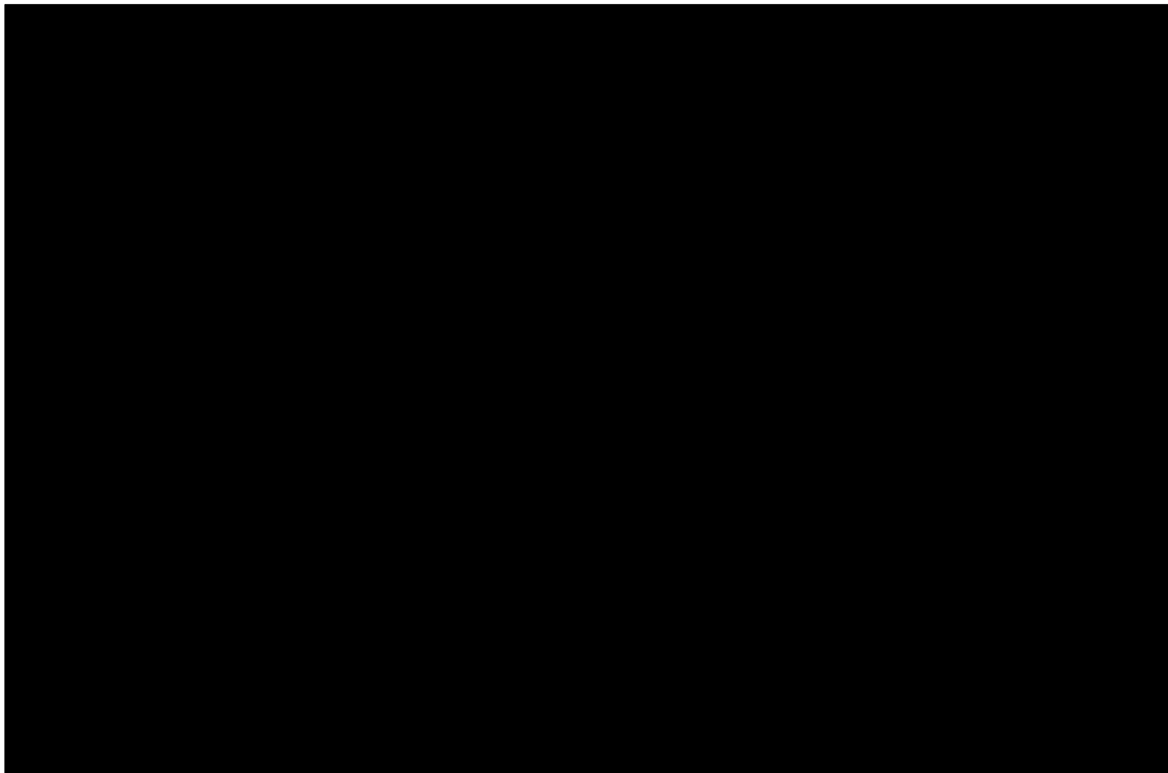
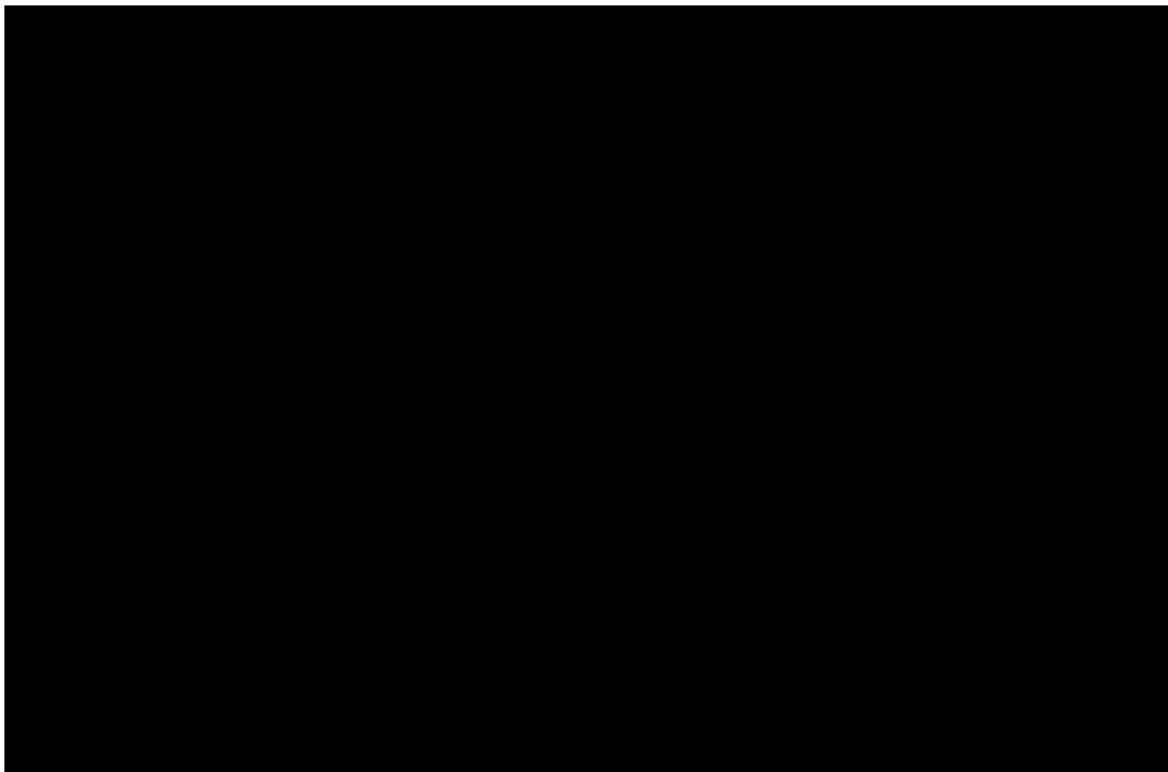


Figure 2: Model projections of proportion alive and free of progression to higher risk states (EAG revised base case)



Following the initial company response to the consultation, the model revisions resulted in an ongoing probability of improvement from the intermediate-low to low risk state beyond 24 weeks on Sotatercept+PDE5i+ERA. It was previously assumed that improvements in risk status following initiation of sotatercept were restricted to occur by 24 weeks. In response to a clarification question on this, the company confirmed that this was due to an error in a formula. They then further revised their base case to allow for improvement in risk status from the intermediate-low to the low risks state for up to 108 weeks, in both the sotatercept and selexipag arm. They justified this based on a further data cut from SOTERIA, which showed risk status improvement or maintenance of relative to the STELLAR week 24 risk status outcome for patients (N=153) who had reached this time point and had available risks status data (see Table 1 of company response document v2). They included the same assumption for selexipag to guard against bias of only including the possibility of longer-term improvement in the sotatercept arm.

The EAG are generally supportive of this change because otherwise the company model does appear to result in projections of risk maintenance for sotatercept that are pessimistic compared with the available data from the SOTERIA extension study. It also helps shift the extrapolated outcomes for selexipag to slightly more optimistic levels.

• [illegible]

To allow for improvement following initiation of prostacyclin analogues, after progression to the intermediate-high risk state on selexipag or sotatercept, the company have revised the structure of their model as outlined in point 7 of their response.

They have added an intermediate-high risk tunnel state, in which patients spend a single cycle of the model following progression to intermediate-high risk, from where they are distributed to a new state reflecting either improvement to intermediate-low risk (with a history of or on PCA), remaining in the intermediate-high risk (with history of or on PCA) or progression to high risk (with history of or on PCA). The chance of achieving improvement is restricted to occur in a single 12 week cycle. This is a pragmatic assumption based on clinical opinion that most patients who improve on prostacyclin analogues do so within 12 weeks of starting. The EAG is satisfied that the approach to restructuring the model is broadly appropriate. The EAG do, however, have concerns with the transition probabilities that are applied in this state.

For the selexipag arm, where the company now assume 85% of patients initiate PCA treatment upon progression (in line with committee preferences), the company apply transition probabilities based on observed transitions from the intermediate-high risk state in the placebo arm of the STELLAR trial between 12 and 24 weeks.

EAG.comment.on.data.to.inform.the.probability.of.improvement.following.PCA.initiation.in. the.selelxipag.arm

This data source seems inappropriate to the EAG, as these transitions reflect those of a cohort that have remained on their background therapy for several months, with the addition of placebo at entry to the trial. Their applicability to a cohort where 85% is newly initiated on PCA following ERS/ESC risk status progression is highly questionable. It suggest only ██ improve their risk status following progression on selexipag, whereas clinical experts consulted by the EAG suggest that 60% could reasonably be expected to shift their risk status back to intermediate-low if initiated on PCA following progression to intermediate-high. The EAG prefers the transition probabilities reported by Roman et al.[4] which were previously used in the company model to reflect the probability of risk status maintenance following initiation of PCA, and are based on data from pivotal PCA trials. These data are, however, based on WHO FC transitions and so may underestimate the chance of ESC/ERS risks status improvement (based on clinical expert advice received by the EAG).

For sotatercept, the company apply a chance of improvement to the 39.9% (in line with committee preference) who are assumed to initiate PCA on top of sotatercept when they progress to intermediate-high risk. For these patients, the company use the observed transitions from the intermediate-high risk state between 12 and 24 weeks in the sotatercept arm of STELLAR. These transitions reflect those of a cohort initiated on treatment with sotatercept 12 weeks prior (many already in the intermediate-high risk state at baseline), rather than a group that have progressed on sotatercept from the intermediate-low risk state, with 39.9% initiating PCA upon progression. It is difficult to assess how relevant these transitions are for the required purpose. Furthermore, for the progressors to intermediate-high risk on sotatercept, the same transitions are used to inform transition probabilities for both those who commence treatment with PCA and those who don't, with the exception that improvements are treated as maintenance of risk status for the 60% who don't initiate PCA analogues.

EAG.comment.on.data.to.inform.the.probability.of.improvement.following.PCA.initiation.in. the.sotatercept.arm

Given the lack of appropriate data to inform the transitions between risk states following initiation of PCA after progression from intermediate-low to intermediate-high risk, there is a uncertainty about the reliability of model output. The EAG retain the company's transition probabilities for those initiating PCA following progression on sotatercept but apply a adjustment that assumes the probability of risk status improvement for initiators of PCA following progression on sotatercept is no worse than it is for those in the selxipag arm. In the long-term, their remains a lower risk of further progression in the history of PCA states in the sotatercept arm compared to the selexipag arm.

• [illegible]

In response to committee's request for revised IV PCA dosing assumptions, the company have accepted that these were above UK averages in their original base case and revised them downwards. They suggest, however, based on feedback from four clinical experts consulted following the first committee meeting, that those in the intermediate-high risk state following previous treatment with selexipag can be suitable for higher target doses than those who have not previously received treatment that targets the prostacyclin receptor (see point 9 of the company response document). Thus, epoprostenol has been costed at a dose of 35ng/kg/minute and Treprostinil at a dose of 70ng/kg/minute in the selexipag arm, compared with doses of 23ng/kg/minute and 42.5ng/kg/minute respectively in the intermediate-high risk state of the sotatercept arm (aligned with committee preferences). The post-selexipag epoprostenol dose, the company state, is based on the midpoint of the range provided by the clinical experts, and the Treprostinil dose is set at twice this level in line with clinical advice. The doses of epoprostenol and Treprostinil in the high risk state of the sotatercept arm are aligned with the higher doses in the selexipag arm.

EAG.comment.on.the.PCA.dosing.assumptions

The EAG has consulted its two clinical experts on these assumptions. Neither were supportive of assuming a higher target dose of IV PCA in those switching from selexipag. They stated that titration may be quicker and better tolerated in those pretreated with prior selexipag exposure but that they would not alter the target dose based on prior/background therapy. Thus, any effect of prior selexipag therapy on expected costs of IV PCA is likely to be small. The EAG, therefore, prefer to retain their original dosing assumptions based on target maintenance doses published in the ESC/ERS guidelines. These are aligned with the doses suggested by the EAG's expert clinical advisors.

In responding to the clarification questions, the company identified an error in their model regarding the sotatercept cost calculations incorporating wastage, whereby the number of 45mg and 60mg vials required for patients in different weight bands were reversed. The company corrected this which results in lower sotatercept acquisition costs. The EAG agree with the correction and include it in their updated analyses.

With the incorporation of clinical benefits for IV PCA now included in the model, through inclusion of a one-off chance of risk status improvement following initiation, the company suggest it is appropriate to include the utility decrement for IV PCA administration, which aligns with clinical and patient expert testimony that was heard at the first committee meeting.

EAG.comment.on.utility.decrement.for.IV.administration.of.PCA

The EAG are in general supportive of this argument but note that the magnitude of the decrement is uncertain, based on methodology that is not directly aligned with the NICE reference case. This was discussed in the original EAG report. The health states for IV PCA use were not reported by patients using a generic descriptive system but rather developed with clinical experts. Whilst the resultant vignettes were valued by a UK general population sample, it is unclear how closely these values would align with the values of EQ-5D health states reported directly by patients with PAH on different modes of treatment administration. The company note insufficient numbers initiating PCA in STELLAR to inform the reduction in EQ-5D values associated with this. The EAG tend to agree with this.

[illegible]

The changes included in the company's revised base case, are summarised in Table 1, with the EAGs comments and suggested changes in columns two and three. The results of the companies revised base case and sensitivity analyses are provided in appendix B of their consultation response document (v2). As outlined in the above critique, the EAG accept several of the company's proposed changes but take issue with others.

The EAG propose a revised preferred base case that:

- Includes the company's revised relative risks, from the MAIC between STELLAR and GRIPHON, to estimate short-term (24 week) transition probabilities for selexipag
- Calibrates the company's new relative risks for application to long-term transition probabilities, so that the model infers a hazard ratio for progression to intermediate-high risk or death that is aligned with the estimated HR for non-fatal clinical progression or death obtained from the company's original ITC between selexipag and sotatercept. This requires 63% of the estimated relative risk reductions to be applied.
- Extends the duration for which there is a chance of improvement on Sotatercept+ PDE5i+ERA or selexipag+PDE5i+ERA, from intermediate-low to low risk, from 24 weeks to 108 weeks.
- Uses the companies revised model structure to allow for a probability of improvement in risk status in the first 12 weeks following initiation of PCA in the intermediate-high risk state.
- Applies the transition probabilities for improvement/deterioration in WHO FC reported by Roman et al.,[4] as proxies for change in risk status following initiation of selexipag in the new intermediate-high risk tunnel state and for ongoing maintenance in the states reflecting a history of PCA use in the selexiag arm. The initial probability of improvement in the sotatercept arm is matched to this in the intermediate-tunnel state, to ensure no determinant in the improvement outcome for those initiating PCA on top of sotatercept+PDE5i+ERA. On an ongoing basis, the source of transition probabilities for the "history of PCA" states in the sotatercept arm assume an increased probability of risk status maintenance over the corresponding states in the selexipag arm, irrespective of PCA use. This is potentially favourable to sotatercept.
- Retains the EAG's IV PCA dosing assumptions, in line with the target doses suggested in the ESC/ERS clinical guidelines, irrespective of health state or background/prior therapy.

- Reinstates the utility decrement for IV PCA administration, with clinical benefits now included in the model through risk status improvement probabilities following PCA initiation in the intermediate-high risk state.

EAG Exploration of further changes to the company model

In reviewing the company's revised model structure, the EAG noted that the calculations for including once-off and ongoing health care resource use associated with PCA initiation and use, were complex and appeared to 1) recount the cost of initiating PCA in each health state in which patients would remain on it; and 2) potentially overestimate ongoing PCA HCRU by applying quarterly costs as an approximation for 12 weekly cycle costs. The modelling of these costs is more complicated in the current iteration of the model due to the chance of initiating PCA in different states in the sotatercept arm and with the use of model averaging. Therefore, the EAG explored a different approach which applied the PCA initiation costs as a once off transition cost in the relevant states of the model, weighted by the proportion of patients expected to initiate PCA in those states. The EAG also assessed the impact of adjusting the ongoing costs per cycle to the exact model cycle length of (12 weeks). Whilst the approach was found to reduce the overall HCRU costs considerably, it affected both arms and so had only a minor impact on the incremental cost and ICER.

Table 1 Summary of company changes, EAG critique and EAG revised base case

Company revision	EAG comment	EAG proposed approach
New relative risks of improvement and deterioration for sotatercept versus selexipag informed by a new MAIC between STELLAR and GRIPHON. Informs short term (24 week) probabilities of improvement and deterioration for selexipag.	The EAG have some concerns regarding the uncertainty around these estimates, related to the reduction in sample size and inability to adjust for all potential effect modifiers. Nevertheless, the EAG accept this as a population adjusted estimate of relative efficacy at 24 weeks.	Retain the company's new relative risks for estimating efficacy out to 24 weeks.
Application of the new relative risks of improvement/deterioration (with no further adjustment) to inform long-term transition probabilities for patients on selexipag.	The EAG have concerns that this approach underestimates the clinical benefits of selexipag, resulting in pessimistic projections of risk status maintenance over time. It also implies that selexipag is no better than placebo at preventing deterioration; the ITC relative risks being based on a measure (WHO FC) for which selexipag demonstrated no benefit in preventing deterioration over placebo in GRIPHON at 26 weeks. There is good evidence, however, for a benefit of selexipag over placebo on clinical progression out to 3 years post-randomisation.	The EAG suggest calibrating the relative risks used to derive long-term transition probabilities for selexipag, so that the model yields weighted average hazards of progression to intermediate-high (as a proxy for clinical worsening) or death, which are consistent with the hazard ratio of ████ that the company derived for the outcome of time to clinical worsening or death in their original ITC.
The chance of improvement on Sotatercept+ PDE5i+ERA or selexipag+PDE5i+ERA, from intermediate-low to low risk, extended from 24 weeks out to 108 weeks.	The EAG acknowledge the uncertainty around this change but accept it because otherwise the company model results in projections of risk maintenance for sotatercept that are pessimistic compared with the available (albeit limited) data from the SOTERIA extension study. The allowance for this in the selexipag arm also guards against bias.	Accept the company's revised approach
Revised model structure that allows for the probability of improvement in risk status in the	The EAG are supportive of the company's structural changes but have issues with the	The EAG prefer to use data from Roman et al. to inform the transition probabilities for those

first 12 week cycle following initiation of PCA. The transition probabilities for sotatercept and selexipag are informed by observed transitions from between 12 and 24 weeks in the sotatercept and placebo arm of the STELLAR trial respectively.	source used to inform transition probabilities of improvement/maintenance for those initiating PCA following progression on selexipag. The company based these on observed transitions between 12 and 24 weeks in the placebo arm of STELLAR, which does not reflect a cohort initiating PCA and produces an estimate for a once-off improvement in risk status which is not well aligned with the EAG's clinical expert opinion. Similarly, the efficacy of adding PCA to sotatercept+PDE5i+ERA is not well informed by the sotatercept arm of STELLAR, but the issue is perhaps less important here since only 39% of patients are assumed to commence PCA in the I-H state of the sotatercept arm. But there is potentially a favourable bias here, with the transitions from the I-H state in the sotatercept arm reflecting a those of a cohort recently initiated on sotatercept rather than placebo.	commencing PCA in the I-H state in the selexipag arm. These transitions between WHO FC states (I-IV) were derived from data reported in pivotal clinical trials of prostacyclin versus placebo, and are used to proxy transitions between ESC/ERS risk states. Th EAG apply a further fix to ensure that the initial probability of improvement is no worse among those commencing PCA on top of sotatercept.
Revised dosing assumptions for IV PCA in the intermediate-high risk and high risk states, these are assumed to be: • 35ng/kg/minute and 70ng/kg/minute for epoprostenol and Treprostinil respectively, following progression on selexipag in both the intermediate-high and high risk states. • 23ng/kg/minute and 42.5ng/kg/minute for epoprostenol and Treprostinil respectively (on top of sotatercept), following progression on sotatercept to the intermediate-high risk state	The EAG's clinical expert advisors were not supportive of higher target doses in those initiating PCA in the I-H risk state following discontinuation of selexipag.	The EAG retain the PCA target doses supported by the ESC/ERS clinical guidelines.

• 35ng/kg/minute and 70ng/kg/minute for epoprostenol and Treprostinil respectively (instead of sotaterept), following progression on sotatercept to the high risk state.		
Removal of the off-sotaterept intermediate-low risk state, with redistribution of these patients to the new intermediate-high risk tunnel state.	This is pragmatic simplifying structural assumption, to avoid the requirement for further complexity in the revised model which accounts for PCA history	The EAG accept this assumption.
Application of the EAGs lower utility value for hospitalisation, in line with committee preferences, and removal of the within state effect of sotatercept on hospitalisation risk.	The EAG agree with this revision	The EAG accept/retain this revision.
Retention of the utility decrement for IV PCA administration, with clinical benefits now included in the model through risks status improvement probabilities following initiation.	The EAG acknowledge there is likely to be substantial utility decrement associated with IV PCA administration within risk states, and that it is now appropriate to include one with the model now revised to better reflect the clinical benefits of initiating PCA. However, the magnitude of the decrement is uncertain, as the methods used for its derivation are not consistent with the NICE reference case.	The EAG include the utility decrement for IV PCA administration and provide further sensitivity analysis around its magnitude. As per the company's alternative scenarios.

[illegible]

The cumulative impact of the EAG's changes to the company revised base are presented in Table 2 below. The resulted EAG revised base case is presented in Table 3, with corresponding probabilities results presented in Table 4. Figures 3 and 4 provide the corresponding cost-effectiveness scatter-plot and acceptability curve.

In line with committee requests, Table 5 provides the results of further scenario analyses, including those exploring the impact of applying baseline characteristics in line with the UK PH audit and alternative approaches to modelling mortality. For transparency, figures showing the proportion surviving and alive and free from progression (to intermediate high risk) by treatment arm have been provided for the followiing scenarios: 1) Baseline characteristics based on UK National Audit of Pulmonary Hypertension (figure 5); 2) Overall survival modelled using the Dependent Gompertz model for the four risk strata (Figure 6); 3) 1 and 2 combined, with recalibration of the long-term transition probabilities to yield a weighted hazard ratio for the reduction in risk of progression or death equal to ████ (Figure 7).

Finally, Table 6 replicates the same scenarios but with no wastage costs included for sotatercept.

Table 2 Cumulative impact of EAG changes to the company's revised base case

Preferred assumption	Incremental costs (£)	Incremental QALYs	Cumulative ICER (£/QALY)	Cumulative ICER with severity modifier (£/QALY)
Company revised base case		2.70		
Calibrate relative risks for application to long-term transition probabilities (63% of relative risks reductions)		1.95		
Apply transition probabilities for improvement/deterioration in WHO FC reported by Roman et al, as proxies for change in risk status following initiation PCA after progression to the intermediate-risk state on selexipag		1.78		
Match initial probability of improvement on PCA initiation in the intermediate-high risk state in the sotatercept arm to Roman et al.		1.78		
Retain IV PCA target doses suggested in the ESC/ERS, irrespective of health state or background/prior therapy		1.78		

Alternative model calculations for HCRU costs associated with PCA initiation and ongoing use		1.78		

Table 3 EAG revised base case results

Treatment	Total LYs	Total QALYs	Total costs (£)	Inc. Cost	Inc. QALYs	ICER	Shortfall multiplier	ICER with multiplier
Sotatercept + dual BGT								
Selexipag + dual BGT					1.78			

Table 4 EAG revised probabilistic base case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	ICER with severity modifier (£/QALY)	ICER weighted according to modifier (£/QALY)
Sotatercept	*******	**	****	*******	**	****	******	******	******
Selexipag	*******	**	****	-	-	-	-	-	-

Figure 3 Cost-effectiveness plane scatter plot (1,000 iterations) – Incremental Costs and QALYs for EAG revised base case (severity modifier applied)

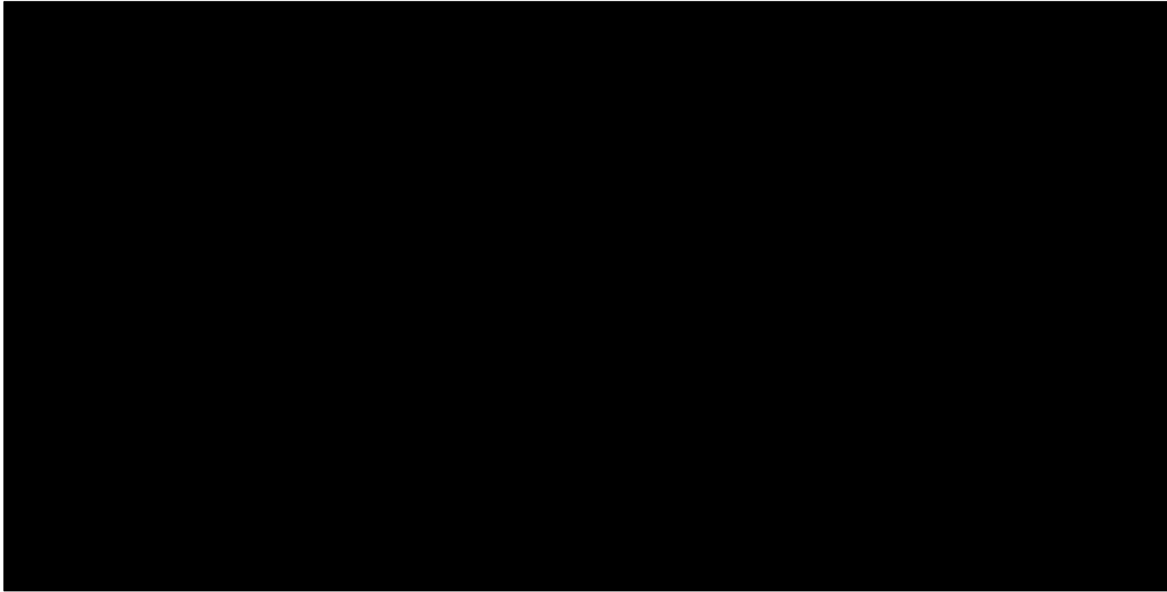


Figure 4 Cost-effectiveness acceptability curve (1,000 iterations) for EAG revised base case (severity modifier applied)

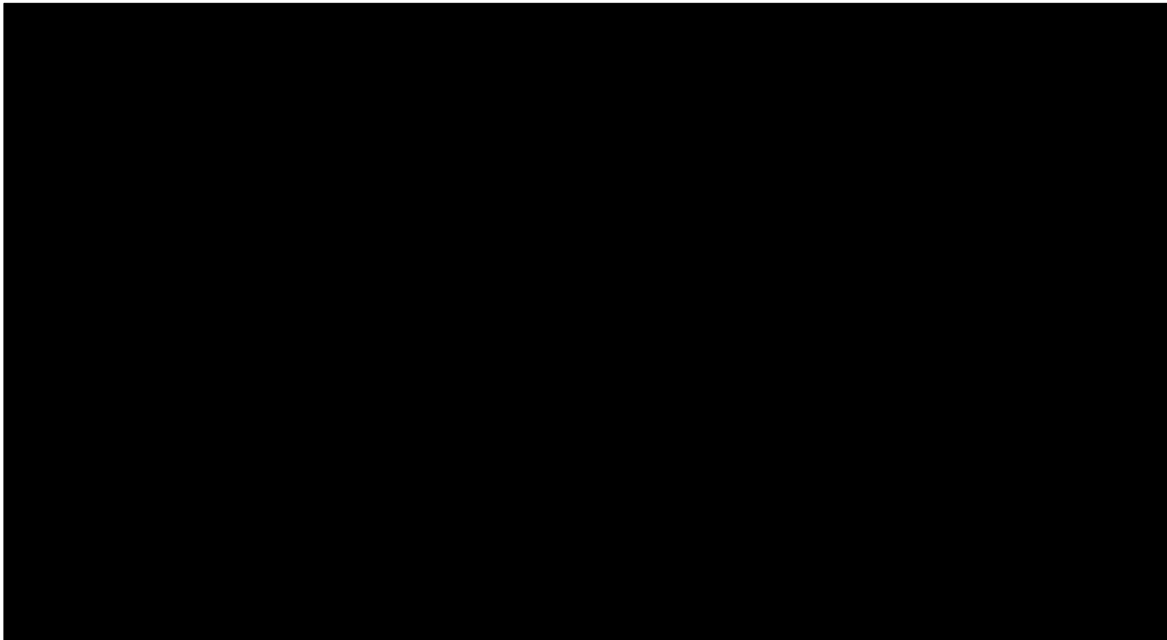


Table 5 Scenario analysis around EAG revised base case

Alternative Scenario	EAG base assumptions	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
EAG base case		████	1.78	████	████
60% probability of risk status improvement following initiation of PCA in the intermediate-high risk state	Once-off Probability of risk status improvement derived from Roman et al.	████	1.74	████	████
Company utility scenario 1	Apply utility decrement of IV PCA administration (= -0.307)	████	1.64	████	████
Company utility scenario 2	Apply utility decrement of IV PCA administration (= -0.307)	████	1.77	████	████
Baseline characteristics based on UK National Audit of Pulmonary Hypertension*	Baseline characteristics of STELLAR	████	1.11	████	████
Overall survival – Dependent Gompertz model for the four risk strata	Single gamma model for the low risk group, with hazard ratios from cox regression applied +	████	2.08	████	████
Overall survival – Dependent Gompertz model for the four risk strata + Baseline characteristics based on UK National Audit of Pulmonary Hypertension*	Baseline characteristics of STELLAR	████	1.69	████	████

	Single gamma model for the low risk group, with hazard ratios from cox regression applied + Baseline characteristics of STELLAR				
Overall survival – Dependent Gompertz model for the four risk strata + Baseline characteristics based on UK National Audit of Pulmonary Hypertension* + recalibrated long-term transition probabilities for selexipag (applying 71% of relative risk reduction)*	Baseline characteristics of STELLAR Single gamma model for the low risk group, with hazard ratios from cox regression applied + Baseline characteristics of STELLAR	████	1.85	████	████

Notes: *Severity multiplier no longer applies in these scenarios

Figure 6 Model projections of proportion alive and free of progression to higher risk states (EAG scenario baseline characteristics based on National PAH Audit)

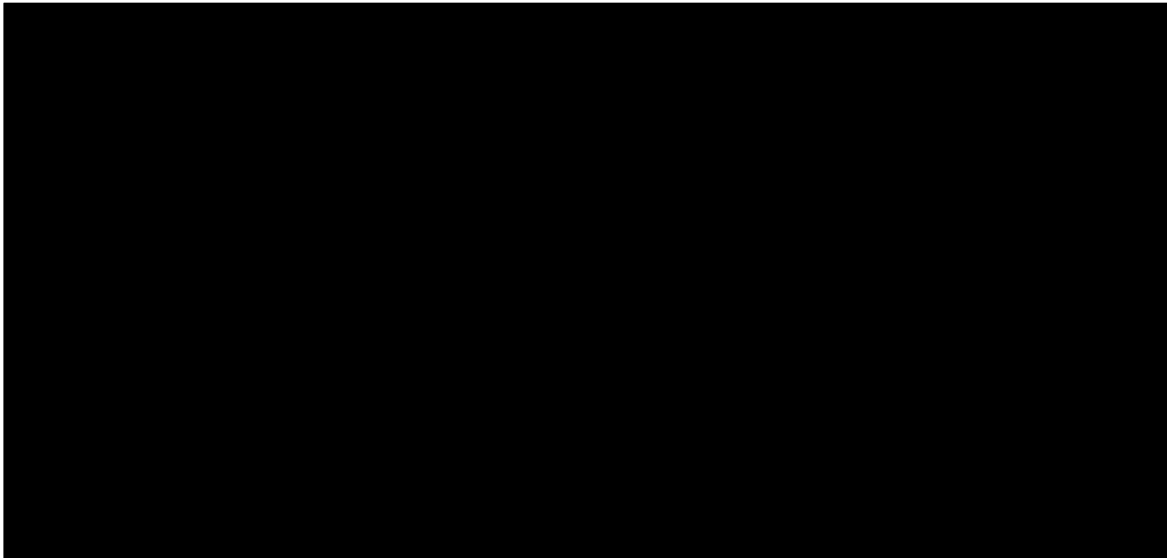


Figure 7 Model projections of proportion alive and free of progression to higher risk states (EAG scenario Mortality based on dependent Gompertz model with for risks strata)

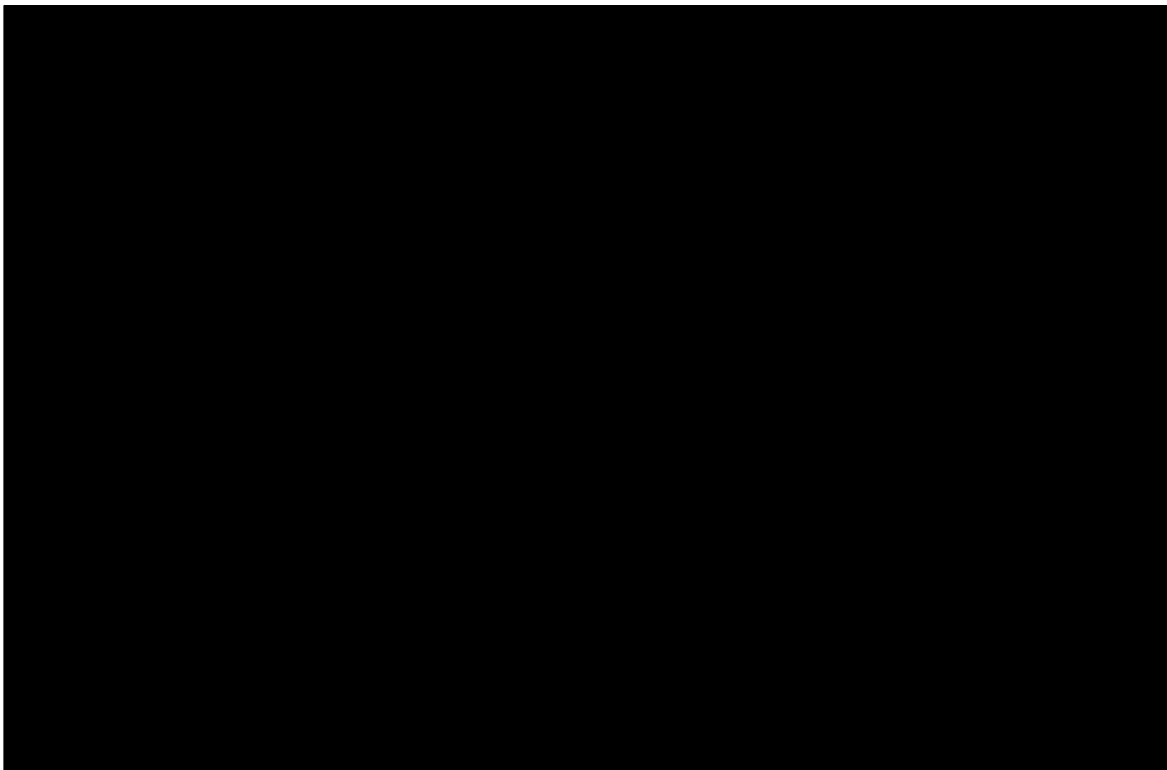


Figure 8 Model projections of proportion alive and free of progression to higher risk states (EAG scenario baseline characteristics based on National PAH Audit + Mortality based on dependent Gompertz model with for risks strata + recalibration of long-term transition probabilities)

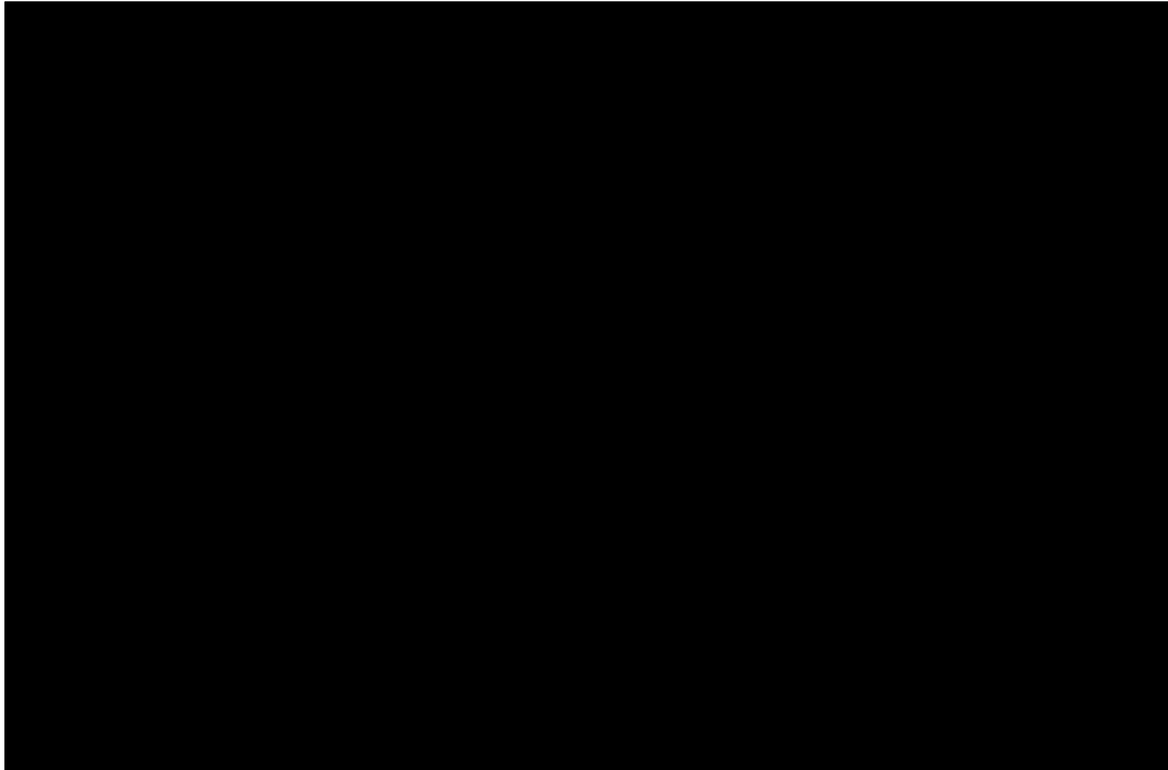


Table 6 Scenario analysis around EAG revised base case (excluding sotatercept wastage costs)

Alternative Scenario	EAG base assumptions	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
EAG base case			1.78		
60% probability of risk status improvement following initiation of PCA in the intermediate-high risk state	Once-off Probability of risk status improvement derived from Roman et al.		1.74		
Company utility scenario 1	Apply utility decrement of IV PCA administration (= -0.307)		1.64		
Company utility scenario 2	Apply utility decrement of IV PCA administration (= -0.307)		1.77		
Baseline characteristics based on UK National Audit of Pulmonary Hypertension*	Baseline characteristics of STELLAR		1.11		
Overall survival – Dependent Gompertz model for the four risk strata	Single gamma model for the low risk group, with hazard ratios from cox regression applied +		2.08		
Overall survival – Dependent Gompertz model for the four risk strata + Baseline characteristics based on UK National Audit of Pulmonary Hypertension*	Baseline characteristics of STELLAR		1.69		

	Single gamma model for the low risk group, with hazard ratios from cox regression applied + Baseline characteristics of STELLAR				
Overall survival – Dependent Gompertz model for the four risk strata + Baseline characteristics based on UK National Audit of Pulmonary Hypertension* + recalibrated long-term transition probabilities for selexipag (applying 71% of relative risk reduction)*	Baseline characteristics of STELLAR Single gamma model for the low risk group, with hazard ratios from cox regression applied + Baseline characteristics of STELLAR		1.85		

Notes: *Severity multiplier no longer applies in these scenarios

• ³⁄Ě³⁄₄¶³⁄₄ ð³⁄₄┛

1. Sitbon O, Channick R, Chin KM, Frey A, Gaine S, Galiè N, Ghofrani HA, Hoeper MM, Lang IM, Preiss R, Rubin LJ, Di Scala L, Tapson V, Adzerikho I, Liu J, Moiseeva O, Zeng X, Simonneau G, McLaughlin VV; GRIPHON Investigators. Selexipag for the Treatment of Pulmonary Arterial Hypertension. N Engl J Med. 2015

2. Howard LS, Rosenkranz S, Frantz RP, Hemnes AR, Pfister T, Hsu Schmitz SF, Skåra H, Humbert M, Preston IR. Assessing Daily Life Physical Activity by Actigraphy in Pulmonary Arterial Hypertension: Insights From the Randomized Controlled Study With Selexipag (TRACE). Chest. 2023 Feb;163(2):407-418

3. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for Population-Adjusted Indirect Comparisons in Health Technology Appraisal. Med Decis Making. 2018 Feb;38(2):200-211

4. Roman A, Barberà JA, Escribano P, Sala ML, Febrer L, Oyagüez I, Sabater E, Casado MA. Cost effectiveness of prostacyclins in pulmonary arterial hypertension. Appl Health Econ Health Policy. 2012 May 1;10(3):175-88.

Table Summary of QALY shortfall analysis

Scenario	Expected total QALYs for the general population	Total QALYs that people living with a condition would be expected to have with current treatment (BAT)	Absolute QALY shortfall	Proportional QALY shortfall	QALY weight
Company revised base case	16.74				1.2
Company revised scenario using baseline characteristics from UK NAPH	13.35				1.2
EAG revised base case	16.74				1.2
EAG revised scenario using baseline characteristics from UK NAPH	13.35				1