NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Final draft guidance

Sotorasib for previously treated KRAS G12C mutation-positive advanced non-small-cell lung cancer

1 Recommendations

1.1 Sotorasib should not be used for treating KRAS G12C mutation-positive locally advanced or metastatic non-small-cell lung cancer (NSCLC) in adults:

- when the cancer has progressed on platinum-based chemotherapy or anti-PD-1/PD-L1 immunotherapy, or
- when these treatments are not tolerated.

1.2 This recommendation is not intended to affect treatment with sotorasib that was funded with managed access before this guidance was published. If this applies, NHS England and the company have an arrangement to make sure people who started treatment during the managed access period will continue to have sotorasib until they and their NHS healthcare professional consider it appropriate to stop.

What this means in practice

These are NICE's final draft recommendations. If these recommendations become final, sotorasib would not be required to be funded and should not be used routinely in the NHS in England for the condition and population in the recommendations.

Issue date: June 2026

Why the committee made these recommendations

This evaluation reviews the evidence for sotorasib for previously treated KRAS G12C mutation-positive advanced NSCLC (NICE technology appraisal guidance 781). It also reviews new evidence collected during the managed access period, which includes evidence from clinical trials and from people having treatment in the NHS in England.

Usual treatment for KRAS G12C mutation-positive locally advanced or metastatic NSCLC in adults when the cancer has progressed on platinum-based chemotherapy or anti-PD-1/PD-L1 immunotherapy, or these treatments are not tolerated, is docetaxel.

Clinical-trial evidence shows that sotorasib increases how long people have before their condition gets worse compared with docetaxel. Clinical-trial and real-world evidence suggest that sotorasib may increase how long people live compared with docetaxel. But it is uncertain how much longer people live with sotorasib compared with docetaxel, because of differences and limitations in each evidence source.

Because of the limitations in the clinical evidence there is uncertainty in the economic model, including how to model the long-term benefits of sotorasib.

When considering the condition's severity and its effect on quality and length of life, the most likely cost-effectiveness estimates are much higher than the range that NICE considers an acceptable use of NHS resources. This is because, based on the committee's preferred assumptions about how to best use the available data, the benefits of sotorasib were small relative to its costs. So, sotorasib should not be used.

2 Information about sotorasib

Marketing authorisation indication

2.1 Sotorasib (Lumykras, Amgen) is indicated for 'the treatment of adult patients with KRAS G12C-mutated locally advanced or metastatic NSCLC, who have progressed on, or are intolerant to, platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy'.

Dosage in the marketing authorisation

2.2 The dosage schedule is available in the summary of product characteristics for sotorasib.

Price

2.3 The list price of sotorasib is £6,907.35 for a 30-day supply of 120 tablets, each containing 240 mg (excluding VAT; BNF online accessed May 2026).

2.4 The company has a commercial arrangement, which would have applied if sotorasib had been recommended.

Sustainability

2.5 Information on the Carbon Reduction Plan for UK carbon emissions for Amgen will be included here when guidance is published.

3 Committee discussion

The evaluation committee considered evidence submitted by Amgen, a review of this submission by the external assessment group (EAG), and responses from stakeholders. See the committee papers for full details of the evidence.

The condition

Details of condition

3.1 Non-small-cell lung cancer (NSCLC) is staged from 1A to 4B according to the size of the tumour, location of involved lymph nodes and the presence of distant metastases. NSCLC diagnosed as stage 3 (locally advanced) or stage 4 (metastatic) is advanced. People with locally advanced NSCLC commonly present with a cough. Other symptoms include shortness of breath, coughing up blood and pain. People with metastatic NSCLC may also have headaches, an enlarged liver, changes in mental health, weakness and seizures. KRAS is a protein that helps control normal cell growth and survival. KRAS is one of the most frequently mutated genes in cancer, including lung cancer. KRAS mutations are linked with a poorer response to treatment and survival outcomes.

In response to draft guidance consultation (from now, referred to as consultation), patient organisations and clinical experts emphasised that people with KRAS G12C mutation-positive NSCLC have a poor prognosis. They noted that this type of cancer is devastating and life-limiting, with many people having rapid disease progression, severe symptoms and a heavy treatment burden. A clinical expert said that, for many people, treatment toxicity and limited tolerability in those with progressed cancer often reduces quality of life and can lead to people avoiding available treatments. A patient organisation also emphasised that having symptoms that are difficult to treat, alongside a lack of targeted treatment options, can be distressing for patients and their carers. The consultation responses all noted that there is an unmet need for effective and tolerable treatments in this population. The NHS England Cancer Drugs Fund (CDF) clinical lead supported this. They emphasised the high demand for targeted treatment options, with around 20 to 40 people seeking access to sotorasib each month through the CDF. The committee acknowledged that advanced NSCLC can substantially affect

health-related quality of life. It concluded that there is an unmet need for targeted treatments to be routinely available in the NHS for KRAS G12C mutation-positive advanced NSCLC.

Clinical management

Comparators

3.2 The company positioned sotorasib as second- and third-line treatment for people with KRAS G12C mutation-positive advanced NSCLC whose cancer has progressed on, or who cannot tolerate, platinum-based chemotherapy or anti-PD-1/PD-L1 immunotherapy, in line with the marketing authorisation. The company's clinical experts advised that many people with NSCLC (without mutations for which there are targeted treatments) now have immunotherapy (alone or with platinum-based chemotherapy) as first-line treatment in the NHS, and few have immunotherapy at second line. So, immunotherapies were not considered relevant comparators. The company explained that platinum-based chemotherapy is rarely used after first-line immunotherapy, so it was also excluded as a comparator.

Sotorasib for previously treated KRAS G12C mutation-positive advanced NSCLC (NICE technology appraisal 781; TA781) included docetaxel and docetaxel plus nintedanib as comparators. The company's clinical experts and the EAG agreed that docetaxel plus nintedanib is no longer frequently used and that docetaxel is more commonly used alone. This is consistent with evidence from the UK Cancer Analysis System (CAS) database. The clinical experts acknowledged there could be differences in treatments used across England. But both confirmed that, in their practices, docetaxel plus nintedanib is used less frequently and use is declining because of toxicities. So, the committee concluded that docetaxel is the only relevant comparator for this evaluation.

Clinical effectiveness

Data sources

3.3 In TA781, the main clinical-effectiveness evidence for sotorasib came from CodeBreaK 100. This was an ongoing phase 1 and 2, single-arm, open-label trial of sotorasib in people with KRAS G12C mutation-positive locally advanced or metastatic NSCLC. The results from an indirect treatment comparison against docetaxel in TA781 suggested a survival benefit for people having sotorasib, although this was uncertain. The committee was aware that a planned phase 3 randomised control trial (CodeBreaK 200) could resolve some of the uncertainty. So sotorasib was recommended for use in the CDF. The key clinical evidence in the current evaluation came from the CodeBreaK 200 trial and also the England-based Cancer Analysis System Real World Evidence (CAS RWE) study. All studies included people with NSCLC who had had at least 1 treatment.

CodeBreaK 200 is an ongoing phase 3, multicentre, open-label, randomised controlled trial comparing sotorasib with docetaxel. It is being done at 149 centres globally, including the UK. The trial recruited 171 people to the sotorasib arm and 174 people to the docetaxel arm. All participants were KRAS G12C mutation positive. The primary outcome was progression-free survival (PFS). Secondary endpoints include overall survival (OS) and health-related quality of life (HRQoL). In the August 2022 data cut, there was a statistically significant improvement for sotorasib compared with docetaxel for PFS (hazard ratio [HR] 0.66, 95% confidence interval [CI] 0.51 to 0.86). A protocol change resulted in a significant reduction in the sample size of the trial. As a result, the trial was no longer powered to detect a statistically significant difference in OS. After adjusting for all treatment switching (switching that happened as per the trial protocol and off-protocol), a post-hoc 2-stage adjustment analysis showed a non-statistically significant OS HR of 0.82 (95% CI 0.32 to 1.31).

CAS RWE was a comparative, retrospective cohort study that used real-world evidence from the CAS database in England. It included evidence from the Cancer Outcomes and Services Dataset (COSD), the Systemic Anti-Cancer Therapy (SACT) dataset, and the Molecular Diagnostics (MDx) dataset. The study included 394 people who had sotorasib and 1,271 people who had docetaxel. The primary objective of the study was to compare OS in people having sotorasib or docetaxel as second-line or later treatment. The secondary objective was to compare time to next treatment or death (TTNTD). Statistical analyses, including propensity score weighting and standardised mortality ratio weighting, were used to balance and weight baseline characteristics between treatment groups. After these adjustments, there was a statistically significant improvement in TTNTD (HR 0.57, 95% CI 0.49 to 0.66) and OS (HR 0.63, 95% CI 0.54 to 0.75).

The committee concluded that CodeBreaK 200 suggested sotorasib improved PFS. The results for OS are uncertain because people in the trial were able to cross-over from the control arm to have sotorasib. Adjusting for the cross-over shows a numerical but not statistically significant improvement in OS, which would not become less uncertain with longer follow-up data. The committee also concluded that the CAS RWE study suggested improvements in TTNTD and OS, but that there were some inherent limitations in the data that meant the results were uncertain (see section 3.4).

Evidence source used to inform clinical-effectiveness inputs

3.4 The company used the CAS RWE data, which was non-randomised, to inform the clinical-effectiveness inputs in the economic model. These were TTNTD used as a proxy for PFS, OS and time to treatment discontinuation or death (TTDD). The company said that the CAS RWE data should be used in the model instead of CodeBreaK 200 data

because it had a larger sample size (n=757 compared with n=345), a longer follow-up (39 months compared with 24 months) and was done in England, whereas CodeBreaK 200 was done in 149 centres globally (including the UK). The committee recalled that because of protocol changes in CodeBreaK 200, the trial was no longer powered to detect a statistically significant difference in OS. During the trial, 59 people (34%) having docetaxel switched to sotorasib instead. Although the company adjusted the OS results for treatment switching, it said it was concerned the results did not align with those from the CAS RWE study and a US-based real-world evidence study (Flatiron). So, it preferred to use the CAS RWE data.

The EAG noted these points but preferred to use evidence from CodeBreaK 200 to inform the clinical-effectiveness inputs in the economic model. It also noted that using CodeBreaK 200 data enabled alignment of sources in the EAG's original base-case economic model for baseline characteristics, OS, PFS and TTDD. The EAG acknowledged that during recruitment to CAS RWE and through statistical and sensitivity analyses, efforts were made to ensure the populations having each treatment were similar. But CAS RWE was not randomised, so it was at risk of selection bias and confounding that cannot be fully addressed by adjustments. During technical engagement, the company provided a risk-of-bias assessment (ROBINS-1) that showed the study was at serious risk of bias. The EAG noted that the KRAS mutation status of most of the people having docetaxel in CAS RWE was unknown, although the impact of mutation status on treatment effect is unclear. The proportions of people with known and unconfirmed mutation status are considered confidential and cannot be reported here. The mutation status of some people in the sotorasib group was also unconfirmed. But because a KRAS G12C mutation was a requirement for having sotorasib, the committee was reassured that most people having sotorasib in CAS RWE had the mutation.

At the first committee meeting, the committee considered the strengths and limitations of both studies. It noted there was a lack of detailed information about the adjustments for treatment switching in CodeBreaK 200. So, the committee was uncertain if the adjustments were appropriate. It also noted a lack of detail about the propensity score analysis that was done on the CAS RWE data. So, the committee was unable to thoroughly assess if this analysis was suitable or robust.

In response to consultation, the company provided additional information about the adjustments for treatment switching in CodeBreaK 200 and the propensity score analysis on CAS RWE. The EAG agreed that the company's adjustment method was valid and that the covariates were generally balanced after the propensity score analysis. But it noted that there may be residual bias because of the missing data on confounders, such as mutation status. The company acknowledged that the missing KRAS mutation status in the CAS RWE docetaxel cohort was a limitation of the data. But it reiterated that this was unlikely to significantly bias the results, emphasising the data it presented at the first committee meeting (see section 3.9). It also presented evidence from ASCO 2022 showing that OS was similar between people with and without the KRAS G12C mutation who were having an immune checkpoint inhibitor plus chemotherapy. The company also consulted 7 UK clinical experts who said that they expected no difference in survival outcomes by KRAS status in this setting. The EAG acknowledged that CAS RWE may better reflect absolute OS for sotorasib, but advised that potential bias remains in the CAS RWE docetaxel data because key confounders were not adjusted for. So, it still preferred CodeBreaK 200 for estimates of relative treatment effects.

The committee acknowledged the strengths of using the same evidence source to inform all clinical outcomes. But it decided that both CAS RWE

and CodeBreaK 200 have important limitations, despite relevant analyses being done to mitigate these. The committee recalled its concerns with the non-randomised comparison in CAS RWE. These included potential risk of bias from imbalances in prognostic factors such as mutation status and brain metastases, and the potential for broader unobserved differences between treatment arms. The committee also noted the variability in the OS HR estimates from the probabilistic sensitivity analyses, including the inverse probability of treatment weighting analysis used to estimate the average treatment effect. This increases the uncertainty around the robustness of the treatment effect estimates from CAS RWE. The committee acknowledged that additional information provided by the company was informative. However, the committee was concerned that the key uncertainties in the clinical evidence sources remained, because of the inherent limitations in each source and because the results for each outcome differed across each source.

The committee concluded that it would consider the most appropriate evidence source for each clinical efficacy input, taking into account the strength of the evidence for each.

Economic model

Company's modelling approach

3.5 The company's model was in line with the model presented in TA781. The company used a partitioned survival model with 3 health states: progression-free, progressive disease and dead. The committee concluded that the model structure was acceptable for decision making.

Progression-free survival

Modelling approach for PFS at the first committee meeting

3.6 In the company's submission, it used the CAS RWE TTNTD data as a proxy for PFS in the economic model. It applied a jointly-fitted generalised

gamma model for both sotorasib and docetaxel. The EAG preferred to use the PFS data from CodeBreaK 200 (see section 3.4), fitting independent log logistic models for both sotorasib and docetaxel.

At the first meeting, the committee noted that the different extrapolation approaches used by the company and EAG had a small impact on the cost-effectiveness results. The committee had concerns about the company's approach of using TTNTD from CAS RWE as a proxy for PFS. It noted that this was a strong assumption, particularly because the PFS results from CodeBreaK 200 (0.66 [95% CI 0.51 to 0.86]) suggest a smaller relative treatment effect than the TTNTD results from CAS RWE (0.57 [95% CI 0.49 to 0.66]). Also, PFS is likely to be shorter than TTNTD. The committee also noted that a high proportion of people had an unknown mutation status in the docetaxel arm, which may have affected the clinical-effectiveness results (see section 3.4). The committee decided to consider 2 alternative approaches that:

- use the CodeBreaK 200 data to inform PFS. The committee agreed that the PFS results from CodeBreaK 200 were of good quality, because the trial was a randomised controlled trial and powered to detect statistically significant differences. Because it was uncertain if TTNTD was an appropriate proxy for PFS, the committee decided it may be reasonable to use CodeBreaK 200 data to inform PFS in the economic model. The committee also noted there were limited other options that were suitable for decision making, given the limitations with CAS RWE
- use sotorasib TTNTD from CAS RWE as a proxy for PFS, but with an adjustment applied. Analyses comparing PFS and TTNTD from CodeBreaK 200 could help explore the relationship between these outcomes. The sotorasib TTNTD data from CAS RWE could then be adjusted to produce a better estimate of PFS in CAS RWE. Then, the inverse of the relative treatment effect for PFS from CodeBreaK 200

would be applied to the sotorasib baseline curve to get the docetaxel curve. The committee noted that in CodeBreaK 200, some people continued having sotorasib after disease progression. This could impact the true relationship between PFS and TTNTD. But the committee agreed that these analyses would still be informative for decision making.

The committee requested that the company explore these 2 approaches as further analyses to understand the impact on the clinical and cost-effectiveness results. The committee considered that these analyses would help decide the most appropriate way to model PFS.

Modelling approach for PFS at the second committee meeting

3.7 After consultation, both the company's and EAG's base-case approach to model PFS was unchanged from the first committee meeting. The company provided the committee's requested scenario analyses. For the scenario using CodeBreaK 200 data to inform PFS, the company chose to model sotorasib using an independent exponential model and docetaxel using an independent log logistic model. But it noted that the results are largely insensitive to the choice of extrapolation because of the maturity of the PFS data. For the scenario using TTNTD data from CAS RWE, the company said that it was inappropriate to apply a different HR to 1 arm of a jointly fitted model. It instead fitted an independent log logistic model for the sotorasib arm, noting that the results were also insensitive to the choice of distribution.

The committee remained concerned about the company's approach to use CAS RWE TTNTD data as a proxy for PFS. It thought that using TTNTD likely overestimates PFS and would need adjustment if used as a proxy. The committee decided that CodeBreaK 200 is a more reliable evidence source for informing PFS because it has benefits of

randomisation from the trial. It also recalled that the PFS data is unlikely to be affected by crossover in CodeBreaK 200, unlike the OS data. The committee decided that for the sotorasib arm, the EAG's base case independent log logistic model had a better statistical fit than the company's scenario using an exponential model. It also thought that the log logistic model provided long-term PFS estimates that aligned more closely with the EAG's clinical expert estimates. It noted that both the company and EAG preferred a log logistic model for the docetaxel arm. The committee concluded that modelling PFS using CodeBreaK 200 data with independent log logistic models for both sotorasib and docetaxel is appropriate for decision making.

Time to treatment discontinuation or death

3.8 In the company's submission, it preferred to use CAS RWE data to inform TTDD, fitting independent generalised gamma models for both sotorasib and docetaxel arms. The EAG's preferred approach was to inform TTDD by fitting parametric survival curves directly to the observed TTDD data from CodeBreaK 200. The EAG requested this at clarification, but the company did not provide it. Without this analysis, the EAG used time-varying HRs from CodeBreaK 200 to inform TTDD in its base case. The committee asked the company to provide the requested analysis to aid its decision making.

After consultation, the company's base-case approach to modelling TTDD was unchanged. The company provided the requested scenario analysis using CodeBreaK 200 TTDD data and fitted an independent log normal model for sotorasib and an independent Weibull model for docetaxel. The EAG updated its base case to use the company's TTDD scenario analysis. But it explained that because the company did not provide sufficient justification for its parametric model selection, the EAG could not assess alternative parametric models to determine the most appropriate extrapolation.

The committee decided that, because it preferred PFS to be informed directly using CodeBreaK 200 data, it would be consistent and appropriate to also model TTDD, a closely related outcome, using parametric models fitted to CodeBreaK 200 TTDD data in the economic model. The committee concluded that modelling TTDD using CodeBreaK 200 data and using an independent log normal model for sotorasib and an independent Weibull model for docetaxel is appropriate for decision making.

Overall survival

Modelling approach at the first committee meeting

3.9 In the company's submission, long-term OS was modelled by applying a jointly-fitted generalised gamma model to the CAS RWE OS data for sotorasib and docetaxel. The EAG preferred to use the CodeBreaK 200 OS data, fitting independent gamma models to the sotorasib and docetaxel arms. At the first committee meeting, the clinical experts stated that they expected 3-year and 5-year OS estimates to be similar to those from the company's base case, which used the CAS RWE study. The company and EAG's OS estimates are confidential so they cannot be reported here. The estimates of the EAG's clinical experts for 3-year and 5-year sotorasib OS were 7% and 2%, respectively. For docetaxel these were 5% and 2%, respectively. Compared with the estimates provided by the clinical experts at the first committee meeting, the EAG clinical expert estimates are lower for sotorasib, and similar or higher for docetaxel.

The committee preferred to use the CAS RWE data to inform the results of the sotorasib group. This was because it was based in England, so the results were likely to best reflect outcomes in the NHS, and had longer follow-up than CodeBreaK 200. But the committee had concerns about the results from the docetaxel group in CAS RWE. This was because of

differences in baseline characteristics in CAS RWE, such as KRAS mutation status and the presence of brain metastases, which could be significant predictors of people's prognosis (see section 3.4). The KRAS mutation status was unknown for most people having docetaxel, so the results may not be relevant to the population being evaluated. The company acknowledged this limitation but explained this was unlikely to bias the results. It explained that KRAS mutation status was unknown because the study included results before sotorasib was available, when routine genetic testing for this mutation was not done. Docetaxel is a non-targeted therapy, so people with a range of mutations can have it. Also, the clinical experts explained that after sotorasib was made available through the managed access agreement, if testing had confirmed the presence of a KRAS G12C mutation, those people would likely have been offered sotorasib. This means that the docetaxel group was more likely to include more people without a KRAS G12C mutation. The company provided a Kaplan–Meier plot showing the OS results of people who had docetaxel, separated by KRAS mutation status. The company suggested this analysis showed that KRAS mutation status may not impact how effective docetaxel is, so the results from the docetaxel group are still relevant. The EAG agreed but noted the sample sizes of people by mutation status were small, so the results of this analysis were unclear.

The committee decided that the CAS RWE data best reflected the absolute sotorasib OS benefits, but not the relative benefit compared with docetaxel. The committee noted that although there were limitations to the OS results in CodeBreaK 200, the relative treatment effect adjusted for treatment switching from CodeBreaK 200 may be a more plausible relative estimate than that from CAS RWE. This was because CodeBreaK 200 was a randomised trial, directly comparing sotorasib with docetaxel, and included the population of interest. The committee also recalled the uncertainties in the docetaxel group of CAS RWE. So, the committee decided that exploring an approach using the CAS RWE data

to inform the baseline curves for sotorasib and then applying the inverse of the relative treatment effect from CodeBreaK 200 would help with decision making.

Modelling approach at the second committee meeting

3.10 In response to consultation, the company updated its base-case approach for estimating long-term OS by applying independently-fitted log logistic models to the CAS RWE OS data for both the sotorasib and docetaxel arms. The company explained that it still preferred to use CAS RWE to inform both absolute survival and relative efficacy, to maintain internal consistency within the model and avoid reliance on the CodeBreaK 200 crossover-adjusted OS estimates. It reiterated that the OS estimates from CodeBreaK 200 were unreliable, both before and after crossover adjustment. The company provided the committee's requested scenario analyses. For the sotorasib arm, it fitted an independent generalised gamma model to the CAS RWE OS data for sotorasib. For the docetaxel arm, it applied the inverse crossover-adjusted relative treatment effect from CodeBreaK 200 to the baseline CAS RWE sotorasib OS curve. But the company stated that this approach increases structural uncertainty in the OS estimates, because it assumes comparability across substantially different populations and settings and relies on strong assumptions.

The committee questioned why the company had changed its preferred approach from the jointly-fitted generalised gamma model to independently-fitted log logistic models. The company explained that, to implement the committee's preferred scenario analysis, it would be methodologically inconsistent to use the sotorasib arm from a jointly-fitted model while applying a different treatment effect. So, it fitted an independent model for the sotorasib OS arm from CAS RWE. This meant that only 1 input parameter needed to change from the company's updated base case to reflect the committee's requested scenario. The company said that independently-fitted log logistic models best captured

the observed hazard patterns, while allowing the use of the inverse crossover-adjusted relative treatment effect for the docetaxel arm. But it noted that both the log logistic and generalised gamma models provided a good statistical and visual fit to the data. The committee remained unclear why the company had moved from a jointly-fitted to an independently-fitted modelling approach. It noted that the company could have retained the jointly-fitted model for the base case, with independent models used for the scenario analysis.

The EAG acknowledged that the company's log logistic model fitted the data well, but thought that it overestimated long-term survival for sotorasib. The EAG preferred the CodeBreaK 200 OS data to inform relative treatment effects because of the risks of confounding in the observational CAS RWE data, including unknown KRAS mutation status and the presence of brain metastases. So, the EAG updated its base case in line with the company's scenario analysis. But it fitted an independent generalised gamma model for the sotorasib arm, instead of an independent log logistic model. It said that this better aligned with the clinical experts' OS estimates from the first committee meeting.

At the second committee meeting, the clinical experts stated that they would expect the 5-year OS estimate for sotorasib to be broadly in line with the estimate generated using the company's new base case. The company and EAG's OS estimates are confidential and cannot be reported here. The committee noted that these OS estimates were higher than those the clinical experts estimated in the first committee meeting, which had previously been aligned with the company's original base case (a jointly-fitted generalised gamma model). For the docetaxel arm, the clinical experts in the committee meeting advised that a 5-year OS estimate of 2% was reasonable and aligned with the EAG's clinical expert estimates.

The committee considered the different modelling approaches presented. It noted that the scenario analysis introduces inconsistency by using CAS RWE to inform sotorasib OS and CodeBreaK 200 to inform docetaxel OS, and does not align with the use of CodeBreaK 200 data for PFS and TTDD. But the committee recalled that the estimates for efficacy outcomes differed between each data source (see section 3.3) and it needs to choose the data source that allows use of best available evidence to inform OS for sotorasib and docetaxel. It decided that CAS RWE may reflect the absolute OS benefit of sotorasib better than CodeBreaK 200, noting that the OS data for sotorasib from CAS RWE was more generalisable to the NHS population. The committee considered the appropriateness of mixing data sources to allow use of best available data to inform the OS extrapolation. It was reassured that the OS estimates for sotorasib were broadly consistent between CodeBreaK 200 and CAS RWE. So, the committee decided that fewer assumptions were needed than would typically be expected when combining data sources. The EAG highlighted that the similar OS value between evidence sources is because the people recruited to the trial and CAS RWE were similar. The committee also acknowledged the high uncertainty associated with the docetaxel arm in CAS RWE because of the methodological limitations of the observational data (see section 3.4). To overcome these methodological limitations in CAS RWE, the committee decided that CAS RWE should be used to estimate absolute OS for sotorasib and CodeBreaK 200 should be used as a more reliable estimate of the relative treatment effect for sotorasib compared with docetaxel. This is in line with the NICE real-world evidence framework. This framework acknowledges that randomised controlled trials are the preferred study design for estimating the causal effects of interventions. However, it suggests using real-world data for baseline rates of events, especially when the data is representative of the target population. So, applying the inverse crossover-adjusted relative treatment effect from CodeBreaK 200 to the sotorasib OS curve derived from CAS RWE was

more appropriate than extrapolating CAS RWE data to inform the docetaxel OS curve.

The committee then considered the following approaches to modelling the sotorasib OS curve using the CAS RWE data:

- the company's independently-fitted log logistic model
- the company's jointly-fitted generalised gamma model (from the first appraisal committee meeting)
- the EAG's independently-fitted generalised gamma model.

The committee decided that all 3 OS curves for sotorasib were plausible. It noted that the EAG's independently-fitted generalised gamma model was the most pessimistic, and the company's independently-fitted log logistic model was the most optimistic. The committee also noted that the company's jointly-fitted generalised gamma model estimates were considered most realistic by the clinical experts at the first committee meeting, and were more optimistic than the EAG's base case and EAG's clinical expert estimates. The committee recalled that the company had not provided a clear justification for changing its base-case model for sotorasib. The EAG had also advised that the company's updated base case overestimated long-term survival for sotorasib, according to clinical expert opinion at the first committee meeting and the EAGs clinical experts. The committee decided that the jointly-fitted model could reduce the risk of overfitting and maximise the use of best available data. It also noted that the 5-year OS estimates for docetaxel were most aligned with those considered the most plausible by the EAG's clinical experts. So, the committee concluded that its preferred approach for modelling long-term OS was:

- using CAS RWE data with a jointly fitted generalised gamma model for sotorasib

- applying the inverse crossover-adjusted relative treatment effect from CodeBreaK 200 to the sotorasib OS curve for docetaxel.

Treatment effect waning

3.11 In TA781, the committee concluded it was plausible that there would be a waning of the relative treatment effect between sotorasib and docetaxel, between 3- and 5- years after starting treatment. For the current evaluation, the company's submission did not include an additional explicit waning assumption in the economic model. It said this was because the follow-up duration for CAS RWE was long enough to implicitly capture any treatment effect waning. The EAG stated that this was a strong assumption, especially because the log-cumulative and smoothed hazard plots for OS (from CodeBreaK 200) showed converging curves, which may suggest waning. The EAG preferred to incorporate an explicit waning assumption, in which the relative OS benefit gradually waned between 2 and 5 years after starting treatment. The committee noted at the first committee meeting that the method of modelling treatment effect waning had very little impact on the cost-effectiveness estimates.

At the second committee meeting, the clinical experts advised that generally there is no treatment effect waning associated with targeted treatments if the treatment benefit is maintained beyond 2 years. They emphasised that a loss of benefit with targeted therapies is often seen in the first 2 years of treatment. The clinical experts also said that the mechanism of how the KRAS mutation is treated by sotorasib explains why there would not be any waning after 2 years. The committee noted that the CodeBreaK 200 data suggests convergence in OS HR between sotorasib and docetaxel, and a similar trend is observed when extrapolating OS from the CAS RWE data. The company explained that the uncertainty in the two-stage crossover adjustment analyses in CodeBreaK 200 makes the OS results difficult to interpret. It also explained that in CAS RWE the hazard in the docetaxel arm slowly

decreases over time, and the log logistic OS model best reflects this decrease.

The committee decided it would be reasonable to apply treatment effect waning if OS is estimated using the inverse HR approach, but not if OS is estimated directly from the clinical data. This is because if OS is estimated directly from the clinical trial then treatment waning is already included. The committee recalled its conclusion of applying the inverse crossover-adjusted relative treatment effect from CodeBreaK 200 to the sotorasib OS curve from CAS RWE data to inform the docetaxel OS curve (see section 3.9). It also noted that treatment effect waning had a small impact on the cost-effectiveness results. The committee concluded that including relative treatment effect waning of sotorasib between 2 to 5 years after treatment initiation is appropriate, in line with the evidence presented to it during this appraisal.

Utility values

Source of utility values and mapping function

3.12 The company used HRQoL data from the EQ-5D-5L questionnaire from CodeBreaK 200. The results from the questionnaire were mapped onto the EQ-5D-3L to align with the NICE reference case. But this was mapped using the van Hout mapping function, rather than the Hernandez-Alava mapping function as stipulated in the NICE reference case. The company preferred using the van Hout mapping function because the Hernandez-Alava function produced higher utilities, which the company thought to be inflated. The company stated this was an artefact of the Hernandez-Alava function and that this was consistent with findings reported in a study by Maervoet and Bergemann (2025). The EAG preferred to use the Hernandez-Alava function in its base case because this aligned with the NICE reference case. Using the Hernandez-Alava mapping function had a small impact on the incremental cost-effectiveness ratio (ICER). The

committee concluded it preferred the Hernandez-Alava mapping function and, in response to consultation, the company used the Hernandez-Alava mapping function.

Approach to modelling utility values at the first committee meeting

3.13 The company's submission estimated utility values by descriptively analysing the HRQoL data by grouped time periods. This was used to inform its time-to-death (TTD) approach. The EAG preferred a progression-based approach to estimate utilities in line with the health states in the model. The company explored different mixed effects models for repeated measures (MMRM) to model utility. But the company did not provide an MMRM that included both progression and TTD covariates, as requested by the EAG. Instead, the company used a continuous TTD approach. Without an MMRM the EAG adopted the company's approach to estimating utilities in the base case, with the exception of using the Hernandez-Alava mapping function (as discussed in section 3.12). Utilities in the EAG's and company's base cases were capped at UK general population values, adjusted for age and sex.

The company stated that it had considered the EAG's request. But to implement the TTD approach fully it would have needed to explore all covariates, which was complex. It also noted that this could not be incorporated in a cohort state model and that doing so would introduce more uncertainty than it resolved. At the first meeting, the committee acknowledged the company's concerns but noted that providing an MMRM would still be useful to explore the interaction between treatment and progression status.

The company used different utility values depending on the treatment people had. The company explained this was because people having sotorasib have fewer treatment-related adverse events and the administration method is less burdensome than for docetaxel. The EAG

advised this was sufficiently justified. This was because the MMRMs that included the treatment covariate had better statistical fit. The committee agreed it was appropriate to use treatment-dependent utilities while people were progression-free. But it noted the company's approach to modelling utilities implied that the lower utility value associated with docetaxel was applied for the whole of the time horizon, even after people stopped having docetaxel. The clinical experts said that the side effects from sotorasib were usually quickly reversible, but the side effects from docetaxel may take weeks to resolve. But they explained that once the side effects had stopped, they would expect the utility for people who had had either treatment to be the same after disease progression. So, the committee asked the company to provide an MMRM with both progression and TTD covariates to help with decision making. The committee also noted that if the TTD approach is used, it should be limited to 6 months before death. In the absence of this MMRM, the committee thought that treatment-independent utilities may also provide a useful estimate.

Approach to modelling utility values at the second committee meeting

3.14 In response to consultation, the company updated its base case using utility values estimated from an MMRM with treatment arm and progression status included as covariates. The company also provided the following additional scenario analyses, using:

- the company's updated base-case utility values including waning of docetaxel post-progression utility decrement (linearly increasing the docetaxel utility value over 12 months to equal sotorasib)
- utility values estimated using an MMRM including both progression status and time-to-death covariates
- a progression-based model with treatment-independent utility values.

The EAG updated its base case in line with the company's first scenario

analysis. It noted that the company's base-case model showed the best statistical fit, and the addition of post-progression docetaxel utility decrement waning aligned with clinical expert advice from the first committee meeting. The EAG explained that the second scenario analysis was suboptimal because of dataset limitations, which reduced the confidence in the estimated utility values. For the third scenario, it advised that treatment-dependent utilities may be appropriate for the PFS health state but this was uncertain after disease progression.

The committee recalled the discussion from the first meeting that once the side effects of docetaxel had stopped, utility for people who had had either treatment was expected to be the same after disease progression. It questioned how long the toxicity with docetaxel persists after stopping treatment. The clinical experts said that most of the side effects of docetaxel would have resolved by 12 months after stopping treatment. The patient expert said that although the physical side-effects of docetaxel may reduce, the mental side-effects may persist, impacting aspects of daily life and overall quality of life. The committee decided that the company's updated base-case utility values are an improvement from the utility values presented at the first committee meeting. But it decided that the EAG's base-case utility values, which capture the temporary side-effects associated with docetaxel after progression, were more accurate. The committee concluded that utility values estimated from the MMRM, with progression status and treatment arm covariates and a 12-month linear waning of docetaxel post-progression utility decrement, are most appropriate.

Costs

3.15 For the costs to administer docetaxel, the company had used the total unit costs. This is an average of the different ways and settings in which docetaxel is administered. The clinical experts explained that docetaxel is usually administered as a day case. So, the committee wanted to see the

economic model updated using day-case unit costs for the administration of docetaxel. In response to consultation, the company adopted the committee's preference for using the day-case unit cost for the administration of docetaxel.

Severity

QALY weighting at the first committee meeting

3.16 The committee considered the severity of the condition (the future health lost by people living with the condition and having standard care in the NHS). The committee may apply a greater weight to quality-adjusted life year gains (QALY; a severity modifier) if technologies are indicated for conditions with a high degree of severity. The company provided absolute and proportional QALY shortfall estimates in line with NICE's technology appraisal and highly specialised technologies guidance manual. The absolute and proportional QALY shortfall estimates were dependent on the evidence source used to inform baseline characteristics and clinical efficacy inputs to calculate QALYs. At the first committee meeting, the company preferred to use CodeBreaK 200 data to inform the baseline characteristics and CAS RWE data to inform the efficacy inputs. The EAG preferred to align the evidence sources, so it used CodeBreaK 200 to inform baseline characteristics and efficacy. The absolute and proportional shortfalls were sensitive to the age used to calculate them:

- in the company's base case, using a mean age of 63 resulted in a proportional shortfall of 0.9510, indicating that a severity weighting of 1.7 should be applied to the QALYs. Using the median age of 64 resulted in a proportional shortfall of 0.9495, indicating a severity weighting of 1.2
- in the EAG's base case, using the median age of 64, absolute shortfall was 10.6 and proportional shortfall was 0.9397, indicating that a severity weighting of 1.2 should be applied.

The committee preferred to use the CAS RWE baseline characteristics for age and sex from the sotorasib arm because they best reflect the population in the NHS. It considered the terminated technology appraisal 1076 (TA1076), which was for the same indication. In TA1076, the company and the EAG applied a severity weighting of 1.7. But the committee noted that, in TA1076, OS data had been immature and so results and QALYs gained were uncertain. So, in this current evaluation, the committee stated that once additional analyses have been provided it would be able to select an appropriate evidence source to inform the clinical-effectiveness inputs (see sections 3.6 to 3.10 and an appropriate approach to modelling utilities (see section 3.13). This would then inform the absolute and proportional shortfalls.

QALY weighting at the second committee meeting

3.17 In response to consultation, both the company and the EAG updated their base case to use CAS RWE to inform the baseline characteristics for age and sex. Using the mean age from the sotorasib arm of CAS RWE resulted in a proportional shortfall of 0.935 in the company's base case and 0.928 in the EAG's base case, corresponding to a severity weighting of 1.2 in both analyses.

The company highlighted that the shortfall estimates are sensitive to modelling assumptions, including age. It noted that several plausible analyses were close to the proportional shortfall threshold for a severity weighting of 1.7. So, it thought that a reliance on point estimates alone does not capture the range of plausible outcomes. The committee acknowledged that none of the scenarios presented by the company in the second committee meeting met the threshold for a severity weighting of 1.7. The company said that NICE has previously taken a pragmatic approach in its appraisals when shortfall estimates were close to the cut-off for 1.7 severity weighting, but it did not provide the committee with

examples of these appraisals. The company reiterated that adagrasib, in the same population, met the 1.7 severity modifier in TA1076 and there has been no meaningful change in disease context or prognosis since that appraisal. It said this suggests that the difference in severity weights may reflect methodological choices rather than changes in disease burden. So, despite estimating a QALY shortfall consistent with a severity weight of 1.2, the company chose to apply a higher severity weighting of 1.7 in its base case.

The committee decided that shortfall estimates should be based on the standard-care QALYs generated within the model of this appraisal. This is because different modelling approaches were used in TA1076 with different modelling assumptions, including for the standard care arm. The committee also noted differences in the baseline age and proportion of women, which may contribute to the differences in estimated shortfalls. It noted that the proportional shortfall in TA1076 was only marginally above the threshold for 1.7 severity weighting. It remained cautious about accepting this weighting because there was insufficient information to assess the robustness of the shortfall estimate, including the strength and reliability of the underlying evidence considered in TA1076. The committee noted that evidence in TA1076 was from a single-arm international study and used a simulation of surrogate endpoints rather than actual OS data. But the evidence to inform shortfall estimates from CAS RWE in the current appraisal only included people in the NHS. Despite its limitations (see section 3.4), this source is more reflective of the population who would have sotorasib in the NHS and more robust than the single-arm international study in TA1076. So, the committee decided it was not appropriate to use TA1076 as a basis for applying a severity weighting in the current appraisal, given the uncertainty in the shortfall estimate. The committee understood that both the company's and EAG's base case and scenario analyses presented at the second committee meeting consistently resulted in a severity weighting of 1.2. It

also noted that the company did not provide evidence to demonstrate that the estimated proportional shortfall did not adequately capture the severity of this population. So, the committee concluded that a severity weighting of 1.2 applied to incremental QALYs is appropriate.

Cost-effectiveness estimates

Company and EAG cost-effectiveness estimates

3.18 The company's and EAG's base cases differed by the following modelling assumptions:

- the clinical evidence source and method used to model PFS (see sections 3.6 to 3.7)
- the clinical evidence source and method used to model TTDD (see section 3.8)
- the clinical-evidence source and method used to model OS (see section 3.9 and section 3.10)
- the modelling of relative treatment effect waning (see section 3.11)
- the approach to modelling post-progression utility values for docetaxel (see section 3.13 and section 3.14)
- the choice of severity weight applied to incremental QALYs (see section 3.16 and section 3.17).

The company's probabilistic base-case ICER is £35,217 per QALY gained (using a severity modifier of 1.7), and the EAG's probabilistic base-case ICER is £102,519 per QALY gained (using a severity modifier of 1.2). When applying a severity modifier of 1.2 in the company's base case, the probabilistic ICER is £49,891 per QALY gained. When applying a severity modifier of 1.7 in the EAG's base case, the probabilistic ICER is £72,367 per QALY gained.

Committee's preferred assumptions

3.19 The committee's preference is to maximise use of the best available data, taking into account the limitations of each data source and its alignment to clinical expert opinion. So its preferred assumptions are to:

- model PFS using CodeBreaK 200 data, and use independent log logistic models for both sotorasib and docetaxel (see section 3.7)
- model TTDD using CodeBreaK 200 data, and use an independent log normal model for sotorasib and an independent Weibull model for docetaxel (see section 3.8)
- model OS using CAS RWE data and a jointly-fitted generalised gamma model for sotorasib, applying an inverse crossover-adjusted relative treatment effect from CodeBreaK 200 to the sotorasib OS curve for docetaxel (see section 3.10)
- include relative treatment effect waning of sotorasib between 2 to 5 years after treatment initiation (see section 3.11)
- using the Hernandez-Alava mapping function to map HRQoL data from EQ-5D-5L to EQ-5D-3L (see section 3.12)
- estimate utility values from an MMRM including progression status and treatment-arm covariates, and include the 12-month linear waning of docetaxel post-progression utility decrement (see section 3.14)
- use day-case unit costs for the administration of docetaxel (see section 3.15)
- apply a severity weighting of 1.2 (see section 3.17).

Using the committee's preferred assumptions, the probabilistic ICER is £108,098 per QALY gained and the deterministic ICER is £102,137 per QALY gained.

Uncaptured benefits

3.20 The committee considered whether there were any uncaptured benefits of sotorasib not fully reflected in the economic model. The patient

organisations, patient experts and clinical experts had emphasised that as an oral treatment taken at home, sotorasib has advantages over docetaxel. It also has a better toxicity profile and greater disease control. They also emphasised that sotorasib is the first targeted treatment for people with previously treated KRAS G12C mutation-positive advanced NSCLC. A patient expert had also said that some adverse effects, including the psychological impact of treatment, may persist for over 12 months after stopping docetaxel (see section 3.14). The committee recalled that a patient expert emphasised the impact to carers, but the committee was not provided with any quantitative evidence to demonstrate this. The committee acknowledged the high unmet need for a targeted treatment (see section 3.1) and concluded that some benefits may not be fully captured in the model.

Acceptable ICER

3.21 NICE's technology appraisal and highly specialised technologies guidance manual notes that, above a most plausible incremental cost-effectiveness ratio (ICER) of £25,000 per QALY gained, judgements about the acceptability of a technology as an effective use of NHS resources will take into account the degree of certainty around the ICER. The committee will be more cautious about recommending a technology if it is less certain about the ICERs presented. But it will also take into account other aspects including uncaptured health benefits. The committee noted that the uncertainty is primarily generated by the:

- limitations in the clinical evidence sources to inform sotorasib and docetaxel OS, PFS and TTDD
- long-term OS estimates for sotorasib and docetaxel used in the economic model.

The committee noted that these issues were drivers of the cost-effectiveness estimates. It recalled that it had chosen a more conservative approach to modelling sotorasib OS than the company's updated base

case. However, its preferred approach was considered plausible at the first committee meeting by the clinical experts and the company, and generates more optimistic estimates than the estimates from the EAG's clinical experts and EAG's base case (see sections 3.9 and 3.10). The committee also recognised that it had mixed data sources to inform the OS extrapolation. The committee considered this to be appropriate, given the limitations in the data sources. Specifically, it decided that the CAS RWE was useful for establishing absolute OS outcomes but less appropriate for estimating relative OS outcomes (see section 3.10).

The committee recalled there are likely to be uncaptured benefits of sotorasib, and noted that it would consider these when determining the most acceptable cost-effectiveness estimate. The committee concluded that an acceptable ICER would be around £35,000 per QALY gained, at the upper end of the cost-effectiveness range.

Risk of decision error

3.22 Section 6.2.32 of NICE's technology appraisal and highly specialised technologies guidance manual states that 'when considering uncertainty, the committee should take into account the likelihood of decision error and its consequences for patients and the NHS'. So, the committee considered the potential benefits and risks to people based on the level of decision uncertainty and whether this could be mitigated. The committee recalled its preference to use the best available data, considering its robustness and consistency with clinical expert opinion. It concluded that the likelihood of decision error has been appropriately considered and was sufficiently low.

Equality

3.23 At the first committee meeting, the patient experts said that some people may have cognitive impairments or disabilities that mean they struggle to self-administer sotorasib. Disability is protected under the Equality Act

2010. But the committee agreed this was not something that could be addressed in its recommendation. The committee acknowledged its duties under the Equality Act 2010. It identified another factor that should also be considered: not everyone is tested for KRAS mutations, so there could be inequalities in access to sotorasib across the NHS. But it decided that because its recommendation does not restrict access to treatment for some people over others, this was not an equalities issue it could address in its recommendation.

In response to consultation, a patient organisation reiterated that inconsistent KRAS testing may already create inequalities in access to targeted treatments. It said that removing access to sotorasib would risk worsening this inequality by limiting options, even when a KRAS mutation is identified. The patient organisation also emphasised that even if recommendations apply equally to all people, its impact is not experienced equally. At the second meeting, the committee acknowledged these concerns but noted that the issues raised related to broader service delivery and diagnostic pathways rather than the evaluation of sotorasib. The committee also considered if the impact of its recommendations may differ across different patient groups. But the committee noted that its recommendation does not restrict access to treatment for some people over others. So, the committee agreed that there were no equality issues it could address in this evaluation. It then considered whether any of the issues raised were health inequality issues. The committee noted that although these concerns were plausible, no robust qualitative or quantitative evidence had been provided to demonstrate that this evaluation would directly worsen or address any health inequalities. So, it could not consider how potential health inequalities could impact its decision making and could not accept a higher level of uncertainty.

Conclusion

Recommendation

3.24 The clinical-effectiveness evidence suggests that sotorasib may improve key outcomes in people with KRAS G12C mutation-positive advanced NSCLC whose disease has progressed on, or who cannot tolerate, platinum-based chemotherapy or anti-PD-1/PD-L1 immunotherapy. The committee considered that its preferred assumptions made the best use of the available data to mitigate against unresolvable limitations in the clinical-effectiveness data. Using these assumptions, the cost-effectiveness estimate is above the range that NICE considers an acceptable use of NHS resources. So, the committee concluded that sotorasib should not be used.

4 Evaluation committee members and NICE project team

Evaluation committee members

The 4 technology appraisal committees are standing advisory committees of NICE. This topic was considered by committee B.

Committee members are asked to declare any interests in the technology being evaluated. If it is considered there is a conflict of interest, the member is excluded from participating further in that evaluation.

The minutes of each evaluation committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Chair

Baljit Singh

Chair, technology appraisal committee B

NICE project team

Each evaluation is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the evaluation), a technical adviser, a project manager, and an associate director or principal technical adviser.

Enna Christmas and Zain Hussain

Technical leads

Alex Filby and Cara Gibbons

Technical advisers

Kate Moore

Project manager

Elizabeth Bell

Principal technical adviser

ISBN: [to be added at publication]