The fate of sotorasib : a regulatory failure potentially harming patients

Ranganathan, Sruthi; Prasad, Vinay; Olivier, Timothée

How to cite

RANGANATHAN, Sruthi, PRASAD, Vinay, OLIVIER, Timothée. The fate of sotorasib : a regulatory failure potentially harming patients. In: Lancet. Oncology, 2024, vol. 25, n° 5, p. 549–552. doi: 10.1016/S1470-2045(23)00616-2

This publication URL: https://archive-ouverte.unige.ch/unige:183971
Publication DOI: 10.1016/S1470-2045(23)00616-2

Title

The fate of sotorasib: a regulatory failure potentially harming patients

Journal

The Lancet Oncology, 25(5)

ISSN

14702045

Authors

Ranganathan, Sruthi
Prasad, Vinay
Olivier, Timothee

Publication Date

2024-05-01

DOI

10.1016/S1470-2045(23)00616-2

Peer reviewed

Quackery

The fate of sotorasib: a regulatory failure potentially harming patients

Introduction

Sotorasib was the first approved KRAS inhibitor granted marketing authorisation by the US Food and Drug administration (FDA) by use of the accelerated approval pathway for patients with advanced non-small-cell lung cancer (NSCLC) with the *KRAS*^{G12C} mutation. The drug targets a subset of mutations in *KRAS*—an oncogene considered undruggable for decades—that occur in 14% of patients with NSCLC.[2]

Sotorasib was approved on the basis of a single-arm, uncontrolled phase 1–2 study (CodeBreaK 100) that reported an objective response rate (ORR) of 37% and a 11·1 month median duration of response. By contrast, other molecular therapies for NSCLC have higher ORRs; osimertinib targets *EGFR*-mutated NSCLC and has an ORR of 90% and alectinib targets *ALK*-mutated NSCLC and has reported an ORR of 82·9%.

The CodeBreaK 200 trial was initially designed as a randomised, post-marketing requirement trial to convert accelerated approval of sotorasib into regular approval. On Oct 5, 2023, the FDA held an Oncology Drug Advisory Committee meeting to discuss whether CodeBreaK 200 could be considered "an adequate and well-controlled trial". The meeting concluded that, on the basis of systemic biases in study conduct, the results of CodeBreaK 200 could not be reliably interpreted.

The manufacturer, Amgen, now faces withdrawal of the drug from the market if it cannot generate a subsequent convincing randomised study. Here, we give an overview of CodeBreaK 200 and address our concerns regarding the design and results, new limitations highlighted by the FDA after accessing the raw study data, the possibility that sotorasib actually reduces survival rather than improves it, and future strategies to facilitate better trial design and conduct.

CodeBreaK 200: what happened?

CodeBreaK 200 was an open-label, randomised, controlled phase 3 trial that recruited patients with a locally advanced, unresectable, or metastatic *KRAS*^{G12C}-mutated NSCLC, who had received platinum chemotherapy and checkpoint inhibitor therapy. Recruited patients were randomly assigned to receive either oral sotorasib or intravenous docetaxel standard of care. The primary endpoint was progression-free survival per blinded independent central review (BICR). Median progression-free survival was 5·6 months (95% CI 4·3–7·8) with sotorasib and 4·5 months (3·0–5·7) with docetaxel, resulting in a significant hazard ratio (HR) of 0·66 (95% CI 0·51–0·86; p=0·0017). The median overall survival was not statistically different between the groups, with a median overall survival of 10·6 months (95% CI 8·9–14·0) in the sotorasib group and 11·3 months (9·0–14·9) in the docetaxel group (HR 1·01 [95% CI 0·77–1·33]; p=0·53).

Initial concerns about the CodeBreaK 200 Trial design and results

On the basis of the initial presentation of the results of CodeBreaK 200 at the 2022 European Society for Medical Oncology (ESMO) congress, we previously raised concerns about the trial design and results. Our five major concerns were use of a suboptimal control group; a change in protocol resulting in inappropriate crossover; a protocol amendment reducing the sample size, thus precluding assessment of survival; limitations in the quality of life (QoL) data; and the potential for informative censoring.

First, the control group in CodeBreaK 200 was suboptimal. At least three other regimens had proven a progression-free survival benefit over docetaxel: paclitaxel plus bevacizumab (IFCT-1130 ULTIMATE trial), docetaxel plus ramucirumab (REVEL trial), and docetaxel plus nintedanib (LUME-Lung 1 trial). These studies showed a progression-free survival benefit, with the REVEL study also reporting an overall survival benefit, and the LUME-Lung 1 study showing an overall survival benefit in a subgroup of patients with adenocarcinoma. Given the reporting of superior outcomes in these studies up to 72 months before patient enrolment started for CodeBreaK 200, these regimens (rather than docetaxel alone) should have been permitted in the control group.

Second, a protocol amendment in CodeBreaK 200 permitted crossover in a trial testing the fundamental efficacy of sotorasib. This allowed patients who initially received docetaxel to receive sotorasib. Since crossover was permitted and no survival benefit was seen, one could argue that patients in both groups benefited from sotorasib, regardless of whether they received the therapy early or late and leading to the non-significant overall survival benefit. However, it is also possible that sotorasib was detrimental, with its harms being masked by the crossover. Elsewhere, we have further discussed the use of crossover in studies seeking to establish efficacy and how problematic and inappropriate this can be.

Department of Medicine, University of Cambridge, Cambridge, UK (SR); **Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA** (VP); **Department of Oncology, Geneva University Hospital, Geneva 1205, Switzerland** (TO)

timothee.olivier@hcuge.ch

SR and TO declare no competing interests. VP reports receiving research funding from Arnold Ventures through a grant made to the University of California San Francisco; royalties for books and writing from Johns Hopkins Press, MedPage, and the Free Press; consulting fees from UnitedHealthcare and OptumRX; receiving revenue from Patreon, YouTube, and Substack for the podcasts Plenary Session, The VPZD Show, and Sensible Medicine; for the newsletters Sensible Medicine, The Drug Development Letter, and VP's Observations and Thoughts; and for the YouTube channel Vinay Prasad MD MPH. This project was funded by Arnold Ventures through a grant paid to the University of California, San Francisco.

For more on the **FDA accelerated approval of sotorasib for *KRAS*^{G12C}-mutated NSCLC** see https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-accelerated-approval-sotorasib-kras-g12c-mutated-nsclc

For more on ***KRAS*^{G12C} mutations in NSCLC** see *N Engl J Med* 2021; **384:** 185–87

For the **CodeBreaK 100 study** see *N Engl J Med* 2021; **384:** 2371–81

For more on **osimertinib in *EGFR*-mutated NSCLC** see *N Engl J Med* 2018; **378:** 113–25

For more on **alectinib in *ALK*-mutated NSCLC** see *N Engl J Med* 2017; **377:** 829–38

For more on the **ODAC meeting on CodeBreaK 200 and relevant meeting documents** see https://www.fda.gov/advisory-committees/advisory-committee-calendar/october-5-2023-meeting-oncologic-drugs-advisory-committee-meeting-announcement-10052023

For more on **CodeBreaK 200** see **Articles** *Lancet* 2023; **401:** 723–46

For the **ESMO 2022 abstract for CodeBreaK 200** see *Annals Oncol* 2022; **33** (suppl 7)**:** s1417–18

For more on **our initial concerns about CodeBreaK 200** see *Trans Oncol* 2023; **28:** 101591

For the **IFCT-1103 ULTIMATE trial** see *Eur J Cancer* 2020; **131:** 27–36

For the **REVEL trial** see **Articles** *Lancet* 2014; **384:** 665–73

For the **LUME-Lung 1 trial** see **Articles** *Lancet Oncol* 2014; **15:** 143–55

For more on the **problems of crossover in randomised trials** see *BMC Cancer* 2023; **23:** 448

For more on **informative censoring** see *Nat Rev Clin Oncol* 2020; **17:** 327–28

Third, a protocol amendment led to a substantial decrease in the sample size. The final sample size of 345 patients, which had decreased from the original sample size of 650, resulted in the study becoming underpowered to detect a significant benefit or a decrease in overall survival outcomes between groups. In fact, the numerical survival data are worse in the sotorasib group than in the control group, with 109 (64%) of 171 patients receiving sotorasib having died as of data cutoff date, compared with 95 (54%) of 174 patients who received docetaxel. On the basis of these data, we calculated that a non-inferiority trial would require 2076 patients to show that sotorasib is non-inferior to docetaxel, with an HR of 1·3, which is substantially higher than the recruited number of patients in CodeBreaK 200.

Fourth, there are limitations in the QoL data in CodeBreaK 200 due to the short timespan (baseline to week 12) over which the QoL of recruited patients was monitored. Considering that the median duration of treatment was 20 weeks with sotorasib, 12 weeks of QoL follow-up would not capture the entire patient experience under this therapy. Also, financial toxicity that might affect real-world patients was not captured in the trial.

Finally, informative censoring might have occurred, meaning that patients censored because of loss to follow-up were different—in terms of characteristics and comorbidities—than those remaining in the study who were censored at data cutoff, and these differences varied by group. A greater proportion of patients in the control group were censored in the first 6 months than patients in the sotorasib group, so the remaining patients in the control group might have been in a better health condition than those remaining in the sotorasib group. Hence the study groups might have become unbalanced in terms of their baseline characteristics, artificially favouring sotorasib over docetaxel. This concern has been discussed elsewhere.[13]

Overall, these limitations led us to believe that not only is there no evidence that sotorasib extends survival, but also that it might even shorten survival compared with an appropriate alternative.

Additional concerns from the FDA analysis

In addition to the concerns highlighted, the FDA raised points regarding the trial design and conduct during the meeting on Oct 5, 2023, more than 2 years after initial approval. Ultimately, the agency concluded that the progression-free survival per BICR could not be reliably interpreted.

The FDA noted there was discordance between the investigator evaluations and BICR evaluations of disease progression and classified this as early or late discordance. Early discordance was when the investigator-assessed progression occurred earlier than the BICR-assessed progression, and late discordance was when the investigator-assessed progression occurred later than the BICR-assessed progression. In CodeBreaK 200, the FDA analysis identified that early discordance occurred more frequently in the docetaxel group than in the sotorasib group, and that late discordance occurred more frequently in the sotorasib group than in the docetaxel group. In other words, investigators tended to delay attributing a progression in the sotorasib group, while tending to hasten labelling a progression in the docetaxel group. This might have been due to investigators being biased to retaining patients on sotorasib for longer, or taking patients off docetaxel earlier. The possibility of a crossover (patients receiving docetaxel crossing to receive sotorasib) might have exacerbated this discordance.

The FDA conducted two sensitivity analyses to show that the progression-free survival benefit of sotorasib over docetaxel could be overestimated under non-extreme assumptions. First, in a tipping point analysis, the FDA analysts made different assumptions regarding the patients in the docetaxel group who dropped out or had early crossover. They found that the HR for the primary endpoint would no longer be significant if these patients had a 50% lower chance of presenting a progression-free survival event than those who remained in the docetaxel group. This assumption was not unlikely. Second, the potential overestimation of the progression-free survival benefit of sotorasib over docetaxel was further supported by an interval-censoring analysis. CT scans were taken every 6 weeks in all patients in both groups to assess for progression of cancer, as defined by prespecified thresholds. Due to the delay between scans, cancer progression events could have occurred any time in the 6 weeks before being detected. By randomising the timing of progression-free survival events over the 6 weeks, the FDA showed that the progression-free survival benefit of sotorasib over docetaxel could be as little as 5 days.

What the FDA did not analyse: the possibility of survival decrement with sotorasib

The possibility of worse overall survival with sotorasib was not explored by the FDA, but must be considered. Given early dropouts, crossover, and the sample size reduction, this possibility remains. In the CodeBreaK 200 Article, in the waterfall plot describing the best response in both groups, early progressors (with an increase in size of baseline tumour burden of >40%) were seen in the sotorasib group, but not in the docetaxel group. Hence, some patients on sotorasib progressed rapidly, whereas this was not seen to as great an extent in the docetaxel

group. Additionally, there was a 10 percentage point higher rate of fatal treatment-emergent adverse events (defined as adverse events that began after the start of trial drugs, regardless of their potential attribution to trial treatment) in the sotorasib group (22%) than in the docetaxel group (12%). It is possible that, without crossover and reduction in the sample size, sotorasib has worse overall survival than docetaxel.

The FDA identified 19 patients who crossed over from docetaxel to sotorasib solely on the basis of investigator-assessed progression-free survival without BICR assessment and estimated they had a median overall survival of 24·4 months. We reconstructed individual patient data from the published Kaplan–Meier survival curves. We identified 19 patients censored in the docetaxel group during the first 2 months, which mirrors the 20 patients that requested withdrawal from the study early after enrolment. For those 19 patients, we modelled the outcomes they would have presented with had they not been censored. First, we assumed that they had the same rate of censoring as the sotorasib group (36% overall). Second, on the basis of 24·4 months median overall survival estimated by the FDA in a subset of patients who crossed over from docetaxel to sotorasib, we assumed that half of the non-censored patients would present with an event over a 24-month period. We randomly allocated hypothetical events over this time period for the 19 patients (sensitivity analysis 1; figure). Finally, we doubled the sample size, nearing the size initially planned for CodeBreaK 200. After running a Cox rank survival analysis, we found an HR of 1·09 (95% CI 0·90–1·32; sensitivity analysis 2; figure). Our findings suggest that, with non-extreme assumptions, a detrimental effect on survival cannot be reasonably ruled out with current data. Even if the confidence intervals cross 1, comparable HRs have resulted in restrictions or withdrawals of drugs in regulatory history—eg, for PARP inhibitors in ovarian cancer.

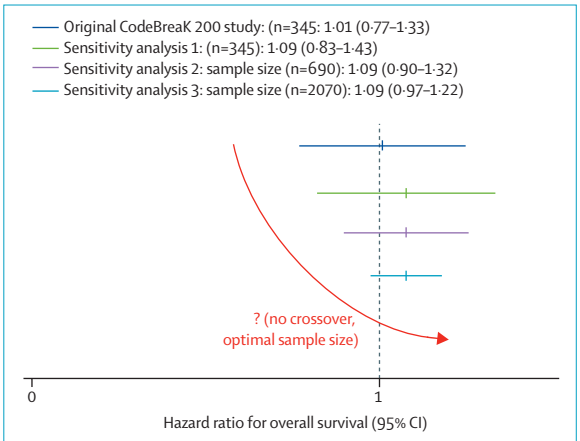

Figure: **Survival sensitivity analyses of CodeBreaK 200 accounting for different events for the 19 patients censored early in the docetaxel group, and with different sample sizes**
We did sensitivity analyses after reconstructing individual patient data based on published Kaplan–Meier curves and made different assumptions for 19 patients who were censored early in the docetaxel group. In sensitivity analysis 1, the sample size was the same as the final sample size in CodeBreaK 200; in sensitivity analysis 2, we simulated outcomes with the sample size that was initially planned for CodeBreaK 200 (n=690); and in sensitivity analysis 3, we used a sample size that we previously estimated would be required if a non-inferiority trial were to be run, on the basis of CodeBreaK 200 results. The red arrow illustrates the trend in hazard ratios and lower 95% CIs over the sensitivity analyses: what would the hazard ratio have been with no crossover, optimal control group, and optimal sample size?

Future measures

What can investigators and sponsors learn from CodeBreaK 200 to improve the design and conduct of future trials? First, trials should aim to be double-blinded or participants and investigators should be educated that newer drugs do not necessarily mean better drugs, and that premature discontinuation is not always beneficial. One of the key contributing factors to the potential systemic bias seen around CodeBreaK 200 was the excitement among patients and investigators that sotorasib could target a previously undruggable molecule. This excitement could have resulted in patients dropping out early from the docetaxel group to receive sotorasib once crossover was permitted, or could have favoured an unconscious bias in investigators evaluating progression in patients. Therefore, measures to educate both participants and investigators regarding equipoise and justifying the trial should be implemented. This systemic bias could also have been solved with a double-blind, double-dummy study design.

Second, the primary endpoints should be ones that matter—in the case of lung cancer, the median overall survival is 11 months, so we do not need to use a surrogate endpoint when survival is this short a duration. Additionally, the FDA showed that crossover and early asymmetric dropout potentially artificially inflated the progression-free survival estimate, showing benefit for those taking sotorasib over docetaxel. Therefore, careful thought should be given to the choice of primary endpoints when considering the trial design.

Third, real-time BICR assessments should be done alongside primary investigator assessments, since the lag between them can make it challenging to investigate discordances, causing informative censoring. If there is a consistent pattern, like the one that occurred in CodeBreaK 200, this could be corrected midtrial.

Fourth, overall survival data for patients who withdrew from the study should be collected to improve the reliability of the survival findings. Given the crossover in CodeBreaK 200, the decrease in the number of participants in the study, and the high dropout rate, it was difficult to adequately answer the question regarding

For more on **reconstructing individual patient data from Kaplan-Meier curves** see *BMC Med Res Methodol* 2021; **21:** 111

For more on **PARP inhibitors and overall survival endpoints in ovarian cancer** see *J Cancer Res Clin Oncol* 2023; **149:** 9509–12

For a **press release on the FDA's new postmarketing requirement** see https://www.amgen.com/newsroom/press-releases/2023/12/amgen-provides-regulatory-update-on-status-of-lumakras-sotorasib

overall survival differences between sotorasib and docetaxel.

Finally, clinical trial data should be shared publicly. The FDA analyses were facilitated by their access to raw individual patient data that were not shared with the public. Promoting data sharing can aid in improving the accuracy of research. Other benefits of data sharing include accelerating clinical research and promoting collaborations.

Conclusion

CodeBreaK 200 has fundamental problems that has made it difficult to make conclusive decisions about the reported results. This was acknowledged by the FDA and we believe it also raises the potential that not only does the drug not extend survival, but it could potentially worsen it. However, despite a controversial trial design and concerning findings, sotorasib remains on the market. Ultimately, the FDA is guilty of regulatory failure: because the agency gave the drug an accelerated approval, it tainted equipoise in the mind of the CodeBreaK 200 investigators, resulting in the use of inappropriate crossover and patients being pulled off of the study differently between the groups. A better regulatory path would have been to approve the drug only after the results of a positive randomised trial.

Now, the FDA turns to a new trial—CodeBreaK 202 (NCT05920356)—to reconsider the conversion approval of sotorasib. But this trial has new problems. It is in the first-line setting, progression-free survival per BICR is the primary endpoint, and it has crossover. Now, it appears unlikely that we will ever adequately answer the clinical question of whether sotorasib improves outcomes for patients beyond what is already achieved with standard of care. Additionally, because on Dec 26, 2023, the FDA issued a new postmarketing requirement for the confirmatory trial to be completed by February, 2028, it will be years before a final regulatory decision on sotorasib is taken, at which point many patients could have been exposed to a potentially toxic, costly, and harmful drug.

*Sruthi Ranganathan, Vinay Prasad, *Timothée Olivier*