FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

ONCOLOGIC DRUGS ADVISORY COMMITTEE (ODAC) MEETING

Virtual Meeting

Thursday, October 5, 2023

9:30 a.m. to 3:03 p.m.

Meeting Roster

ACTING DESIGNATED FEDERAL OFFICER (Non-Voting)

Joyce Frimpong, PharmD

Division of Advisory Committee and

Consultant Management

Office of Executive Programs, CDER, FDA

Mark R. Conaway, PhD

Professor

Division of Translational Research and Applied

Statistics

Department of Public Health Sciences

The University of Virginia School of Medicine

Charlottesville, Virginia

William J. Gradishar, MD

Professor of Medicine/Betsy Bramsen Professor of Breast Oncology

Chief, Hematology/Oncology

Robert H. Lurie Comprehensive Cancer Center

Feinberg School of Medicine at Northwestern University

Chicago, Illinois

Ravi A. Madan, MD

(Chairperson)

Senior Clinician

Head, Prostate Cancer Clinical Research Section

Genitourinary Malignancies Branch

Center for Cancer Research (CRC)

National Cancer Institute (NCI)

National Institutes of Health (NIH)

Bethesda, Maryland

David E. Mitchell

(Consumer Representative)

President

Patients for Affordable Drugs

Bethesda, Maryland

Jorge J. Nieva, MD

Associate Professor of Clinical Medicine

Section Head, Solid Tumors

University of Southern California (USC) Norris

Comprehensive Cancer Center

Keck School of Medicine of USC

Los Angeles, California

Ashley Rosko, MD

Associate Professor

Division of Hematology

Medical Director Oncogeriatric

The Ohio State University Comprehensive

Cancer Center

Columbus, Ohio

Daniel Spratt, MD

Vincent K Smith Chair, Department of Radiation Oncology

Professor of Radiation Oncology and Urology

University Hospitals Seidman Cancer Center

Case Western Reserve University

Cleveland, Ohio

Neil Vasan, MD, PhD

Assistant Professor

Division of Hematology & Oncology

Department of Medicine

Herbert Irving Comprehensive Cancer Center

Columbia University Medical Center

New York, New York

ACTING INDUSTRY REPRESENTATIVE TO THE COMMITTEE

(Non-Voting)

Albert L. Kraus, PhD

(Acting Industry Representative)

Managing Partner

GDS Partners, LLC

Guilford, Connecticut

TEMPORARY MEMBERS (Voting)

James L. Gulley, MD, PhD, FACP

Medical Oncology

Co-Director

Center for Immuno-Oncology, CRC, NCI

Clinical Director, NCI, NIH

Bethesda, Maryland

Philip C. Hoffman, MD

Clinical Professor of Medicine

Section of Hematology/Oncology

University of Chicago

Chicago, Illinois

James Pantelas

(Patient Representative)

Howell, Michigan

Pamela Shaw, PhD, MS

Senior Investigator

Biostatistics Division

Kaiser Permanente Washington Health Research Institute

Seattle, Washington

FDA PARTICIPANTS (Non-Voting)

Richard Pazdur, MD

Director, Oncology Center of Excellence (OCE)

Office of the Commissioner (OC)

Director (Acting)

Office of Oncologic Diseases (OOD)

Office of New Drugs (OND), CDER, FDA

Harpreet Singh, MD

Director

Division of Oncology 2 (DO2)

OOD, OND, CDER, FDA

Paz Vellanki, MD, PhD

Cross Disciplinary Team Lead

DO2, OOD, OND, CDER, FDA

Jeevan Puthiamadathil, MD

Clinical Reviewer

DO2, OOD, OND, CDER, FDA

Pallavi Mishra-Kalyani, PhD

Deputy Director

Division of Biometrics V (DBV)

Office of Biostatistics (OB)

Office of Translation Sciences (OTS)

CDER, FDA

Anup Amatya, PhD

Biometrics Team Leader

DBV, OB, OTS, CDER, FDA

Chi (Chuck) Song, PhD

Statistical Reviewer

DBV, OB, OTS, CDER, FDA

C O N T E N T S

C O N T E N T S (continued)

P R O C E E D I N G S

(9:30 a.m.)

Call to Order

DR. MADAN: Good morning, and welcome. I'd first like to remind everyone to please mute your line when you are not speaking. For media and press, the FDA press contact is Lauren-Jei McCarthy. Her e-mail is currently displayed.

My name is Ravi Madan. I will be chairing this meeting. I will now call the October 5, 2023 Oncologic Drug Advisory Committee meeting to order. Dr. Joyce Frimpong is the acting designated federal officer for this meeting and will begin with introductions.

Introduction of Committee

DR. FRIMPONG: Good morning. My name is Joyce Frimpong, and I'm the acting designated federal officer for this meeting. When I call your name, please introduce yourself by stating your name and affiliation.

Dr. Conaway?

DR. CONAWAY: Mark Conaway, biostatistics,

University of Virginia School of Medicine.

DR. FRIMPONG: Thank you.

Dr. Gradishar?

DR. GRADISHAR: Bill Gradishar, medical oncology, Northwestern University.

DR. FRIMPONG: Dr. Madan?

DR. MADAN: Ravi Madan, medical oncologist, National Cancer Institute.

DR. FRIMPONG: Mr. Mitchell?

MR. MITCHELL: I'm David Mitchell. I am the consumer representative to the ODAC. I am president of an organization called Patients for Affordable Drugs, and I'm a multiple myeloma patient myself.

DR. FRIMPONG: Dr. Nieva?

DR. NIEVA: Hello. I'm George Nieva. I'm a thoracic medical oncologist at the University of Southern California Norris Cancer Center.

DR. FRIMPONG: Dr. Rosko?

DR. ROSKO: Good morning. Ashley Rosko, Division of Hematology at The Ohio State University.

DR. FRIMPONG: Thank you.

Dr. Spratt?

DR. SPRATT: Dr. Dan Spratt. I'm a professor and chair of radiation oncology at Case Western Reserve University.

DR. FRIMPONG: Dr. Vasan?

DR. VASAN: Hi. Neil Vasan, medical oncologist at Columbia University, Irving Medical Center.

DR. FRIMPONG: For our industry representative, Dr. Kraus?

DR. KRAUS: Good morning, everyone. Albert Kraus. I'm an independent consultant with GDS Partners, and prior, a lot of industry experience, small and big, in R&D. Thank you.

DR. FRIMPONG: Our temporary voting members, Dr. Gulley?

DR. GULLEY: Hi. James Gulley, National Cancer Institute, medical oncology.

DR. FRIMPONG: Dr. Hoffman?

DR. HOFFMAN: My name is Philip Hoffman. I'm a medical oncologist at University of Chicago.

DR. FRIMPONG: Mr. Pantelas?

MR. PANTELAS: I am Jim Pantelas. I'm a patient advocate and a lung cancer survivor of 18 years.

DR. FRIMPONG: And Dr. Shaw?

DR. SHAW: Hello. My name is Pamela Shaw, and I'm senior investigator of biostatistics at Kaiser Permanente Washington Health Research Institute.

DR. FRIMPONG: And now for our FDA participants, Dr. Pazdur?

DR. PAZDUR: Hi. Rick Pazdur, director, Oncology Center of Excellence, FDA.

DR. FRIMPONG: Dr. Singh?

DR. SINGH: Harpreet Singh, medical oncologist, director of the Division of Oncology 2.

DR. FRIMPONG: Dr. Vellanki?

DR. VELLANKI: Hi. Paz Vellanki, medical oncologist and cross-disciplinary team lead at the FDA.

DR. FRIMPONG: Dr. Puthiamadathil?

DR. PUTHIAMADATHIL: Hi. Jeevan

Puthiamadathil, medical oncologist and clinical reviewer in the Division of Oncology 2 at the FDA.

DR. FRIMPONG: Dr. Mishra-Kalyani?

DR. MISHRA-KALYANI: Pallavi Mishra-Kalyani, deputy division director, Division of Biometrics V.

DR. FRIMPONG: Dr. Amatya?

DR. AMATYA: Anup Amatya, statistical team leader, Division of Biometrics V. Thank you.

DR. FRIMPONG: And Dr. Song?

DR. SONG: Chuck Song, the primary statistical reviewer from FDA.

DR. FRIMPONG: Thank you, everyone.

Dr. Madan, I'll hand it back to you.

DR. MADAN: Thank you.

For topics such as those being discussed at this meeting, there are often a variety of opinions, some of which are quite strongly held. Our goal at this meeting will be a fair and open forum for discussion of these issues, and that individuals can express their views without interruption. Thus, as a gentle reminder, individuals will be allowed to speak into the

record only if recognized by the chairperson. We look forward to a productive meeting.

In the spirit of the Federal Advisory Committee Act and the Government in the Sunshine Act, we ask the advisory committee members to take care that their conversations about the topic at hand take place in the open forum of the meeting.

We are aware that members of the media are anxious to speak with the FDA about these proceedings; however, the FDA will refrain from discussing the details of this meeting with the media until its conclusion. Also, the committee is reminded to please refrain from discussing the meeting topic during breaks or lunch. Thank you.

Dr. Frimpong will read the Conflict of Interest Statement of the meeting.

Conflict of Interest Statement

DR. FRIMPONG: Thank you, Dr. Madan.

The Food and Drug Administration is convening today's meeting of the Oncologic Drugs Advisory Committee under the authority of the Federal Advisory Committee Act of 1972. With the

exception of the industry representative, all members and temporary voting members of the committee are special government employees or regular federal employees from other agencies and are subject to federal conflict of interest laws and regulations.

The following information on the status of this committee's compliance with federal ethics and conflict of interest laws, covered by but not limited to those found at 18 U.S.C. Section 208, is being provided to participants in today's meeting and to the public.

FDA has determined that members and temporary voting members of this committee are in compliance with federal ethics and conflict of interest laws. Under 18 U.S.C. Section 208, Congress has authorized FDA to grant waivers to special government employees and regular federal employees who have potential financial conflicts when it is determined that the agency's need for a special government employee's services outweighs their potential financial conflict of interest, or

when the interest of a regular federal employee is not so substantial as to be deemed likely to affect the integrity of the services which the government may expect from the employee.

Related to the discussion of today's meeting, members and temporary voting members of this committees have been screened for potential financial conflicts of interests of their own, as well as those imputed to them, including those of their spouses or minor children and, for purposes of 18 U.S.C. Section 208, their employers. These interests may include investments; consulting; expert witness testimony; contracts, grants, CRADAs; teaching, speaking, writing; patents and royalties; and primary employment.

Today, the discussion of supplemental new drug application, sNDA, 214665/S-005, for Lumakras, sotorasib, tablets, submitted by Amgen, Incorporated, for the proposed treatment of adult patients with KRAS G12C mutated locally advanced or metastatic non-small cell lung cancer, as determined by an FDA approved test, who have

received at least one prior systemic therapy. This supplement proposes to convert the NDA to full approval, based on the confirmatory study, CodeBreaK 200. The committee will consider the result of the CodeBreaK 200 study and discuss the benefit-risk profile of Lumakras.

This is a particular matters meeting during which specific matters related to Amgen's sNDA will be discussed. Based on the agenda for today's meeting and all financial interest reported by the standing voting members and temporary voting members, no conflict of interest waivers have been issued in connection with this meeting.

To ensure transparency, we encourage all standing committee members and temporary voting members to disclose any public statements that they have made concerning the product at issue. With respect to FDA's invited industry representative, we would like to disclose that Dr. Albert Kraus is participating in this meeting as a non-voting industry representative, acting on behalf of regulated industry. Dr. Kraus' role at this

meeting is to represent industry in general and not any particular company. Dr. Kraus is employed by GDS Partners, LLC.

We would like to remind members and temporary voting members that if the discussion involves any other products or firms not already on the agenda for which an FDA participant has a personal or imputed financial interest, the participants need to exclude themselves from such involvement, and their exclusion will be noted for the record. FDA encourages all other participants to advise the committees of any financial relationships that they may have with the firm at issue. Thank you.

Dr. Madan, back to you.

DR. MADAN: Thank you, Dr. Frimpong.

We will now proceed with the FDA introductory remarks from Dr. Harpreet Singh.

FDA Opening Remarks - Harpreet Singh

DR. SINGH: Good morning. I am Harpreet Singh, medical oncologist and director of the Division of Oncology 2. We convene today's

Oncologic Drugs Advisory Committee to discuss the development of sotorasib for patients with non-small cell lung cancer harboring KRAS G12C mutations. Sotorasib was granted an accelerated approval based on single-arm data from the trial CodeBreaK 100 in May of 2021, making it the first FDA approved therapy to target KRAS G12C, which for many decades was considered an undruggable target in oncology.

Amgen, who I will refer to as the applicant moving forward, conducted a randomized trial, CodeBreaK 200, to verify the clinical benefit of sotorasib. The applicant submitted a supplemental new drug application in February for conversion from accelerated or conditional approval to full or traditional approval, based on the CodeBreaK 200 results, which will serve as the focus for today's committee.

FDA's initial assessment of CodeBreaK 200 was that the trial was reported as statistically significant, meeting its primary endpoint of progression-free survival, but with a small or

incremental effect against single-agent docetaxel. There was no difference in overall survival between the two treatment arms. However, initial signs of potential bias, such as high rates of patient dropout on the docetaxel relative to the sotorasib arm, led us to further investigate the potential for systemic and open-label bias in CodeBreaK 200. The patterns of behavior and study conduct suggested a consistent bias in favor of sotorasib and created uncertainty in our ability to interpret the results of the primary efficacy endpoint, and in turn, the overall trial.

Sotorasib is an oral tyrosine kinase inhibitor developed for the treatment of patients with KRAS G12C mutations, which comprise approximately 13 percent of patients with non-small cell lung cancer. In May of '21, the results of CodeBreaK 100 led to an accelerated approval of single-agent sotorasib for patients who have progressed after one line of systemic therapy. Most patients had received a standard first-line regimen of immunotherapy with platinum-based

chemotherapy. CodeBreaK 100 was a single-arm trial, yielding a 36 percent response rate with 10 months of durability, for a population with few options often relegated to single-agent docetaxel or other chemotherapies, with historic response rates of 8 to 12 percent.

Sotorasib was a first-in-class therapy, and early promising results were met with great enthusiasm by the oncology community. In today's information age, it is possible that emerging data from other trial results may have increased patients and investigator awareness of sotorasib, and in turn, their desire to access sotorasib, making it more challenging to conduct an open-label trial.

CodeBreaK 200 is an ongoing randomized trial designed to verify the clinical benefit of sotorasib seen in early single-arm data. Patients were randomized 1 to 1 to receive daily oral sotorasib versus every 3-week intravenous docetaxel. The primary endpoint was progression-free survival by blinded independent

central review. Enrollment began in June of 2020 and was completed by April of 2021, prior to the accelerated approval in May the same year.

Crossover was not initially offered as part of the study design; however, with the results of CodeBreaK 200 in hand, FDA and the applicant discussed adding crossover to mitigate concerns for patient and investigator bias in favor of the investigational drug. Though the study design was modified to add crossover to allow patients to progress on docetaxel to access sotorasib, the patterns of early dropout on the docetaxel arm had already occurred, and additional signs of potential bias were beginning to emerge.

Before we discuss the top-line results, we note that the FDA was contacted by the applicant several months before the final CodeBreaK 200 results were submitted to discuss results of a planned interim PFS analysis, which had been narrowly flipped from a negative to a positive finding, based on applicant-triggered re-reads of discrepant assessments between investigators and

blinded radiologists.

While FDA advised that the applicant not submit for regulatory consideration at that time and instead follow the data monitoring committee's advice to continue the trial as planned, this was the first suggestion that the sotorasib arm may have underperformed or docetaxel overperformed relative to historical data. This was also when we first became concerned about possible violations of the imaging charter and overall study integrity. You will hear more about this from Dr. Puthiamadathil later in the FDA presentation.

Top-line results of the primary endpoint, PFS by BICR, and overall survival, which was a secondary endpoint, are shown here. Per the applicant, treatment with sotorasib yielded a statistically significant improvement in PFS relative to docetaxel, with a median improvement of 5 weeks and a hazard ratio of 0.66. We note that patients' tumors were measured every 6 weeks, creating uncertainty in the median PFS benefit of 5 weeks, since tumors could have begun growing

earlier than imaging picked up.

FDA acknowledges the head-to-head design of CodeBreaK 200, as well as the different routes of administration and toxicity profiles of sotorasib versus docetaxel. There was no difference in overall survival, and though up to 34 percent of patients on the docetaxel arm subsequently were treated with KRAS targeting therapies, our statistical review indicates that this was unlikely to have impacted these OS findings. In a refractory disease setting, overall survival is a critical measure of benefit when assessing the totality of evidence.

One of the first signals of potential bias in CodeBreaK 200 was the high rate of early dropouts on the docetaxel arm relative to sotorasib. The FDA observed that 23 patients randomized to the docetaxel arm dropped out of the study or withdrew consent shortly after they were made aware of their treatment assignment, compared to only 2 patients on the sotorasib arm. This high rate of imbalanced early dropout from the docetaxel

arm was a signal that CodeBreaK 200 may have had a perceived lack of equipoise by both patients and providers elect.

When patients drop out of trials or withdraw consent in an asymmetric manner, this results in a loss of information, which could potentially bias results and make it difficult to quantify the true treatment effect. Dropout in clinical trials is common, particularly in the setting of open-label trials. There may be a perceived loss of equipoise if patients and/or providers believe that the control arm is suboptimal. Emerging data from other trial results in an active therapeutic landscape may influence rate of drop out, and while this type of open-label bias would be concerning in any trial, this is compounded in CodeBreaK 200 by the modest effect of sotorasib relative to docetaxel.

In addition to the asymmetric early dropout, FDA found that investigators' assessment of patients' imaging or CT scans often ruled in favor of the sotorasib arm. Such patterns of behavior

uncovered in the data can permeate to other aspects of trial conduct, which we are not able to see or quantify. For example, there may be underreporting of adverse events in an effort to remain on study drug.

The FDA review found that collectively investigator assessments relative to that of the blinded independent central review were biased in favor of sotorasib. When considering discordance assessment between the unblinded investigators and the BICR, we found that there were greater early calls of progression compared to the blinded review for the docetaxel arm, or early discordance, and there were more late calls for progression by investigators compared to the BICR for the sotorasib arm, which is called late discordance. Again, while we do expect some discordance in every trial, in CodeBreaK 200, the differential distribution of discordance across arms may signal a potential bias in favor of sotorasib.

Some of the patterns in study conduct and resulting challenges in interpretation of the data

raised questions as to whether we could consider CodeBreaK 200 an adequate and well-controlled trial. According to the Code of Federal Regulations, features of an adequate and well-controlled trial include adequate measures to minimize bias in subject assignment to treatment group; adequate measures to minimize bias on the part of subjects, observers, and analysts of the data; well-defined and reliable methods to assess response; and ultimately, whether the study allows for adequate analysis of the results to assess the effect of the drug.

In the primary FDA presentation, our oncologists and biostatisticians will describe how the asymmetric early dropout, discordance between investigators and the BICR in assessment of progressive disease, and potential violations of the imaging charter have made it challenging to truly assess the treatment effect of sotorasib in CodeBreaK 200. The loss of patient-level information due to censoring and early dropout confounds our ability to conduct adequate analyses

of the results of CodeBreaK 200 to assess the effect, and importantly the magnitude of effect, of sotorasib versus docetaxel.

You will hear from Dr. Chuck Song, FDA biostatistician. Dr. Chuck Song conducted a tipping-point analysis, examining how the observed results may change if we were to assume a different risk of PFS event for 39 patients, who were censored early due to either dropout or crossover. In this analysis, we show a potential loss of statistical significance in the PFS endpoint, suggesting that the primary endpoint may not be sufficiently robust or able to withstand variability in patient outcome.

A complete and balanced assessment of the primary PFS endpoint includes evaluation of the hazard ratio, median benefit, event rates, and shape of the Kaplan-Meier curves. The applicant asserts that the results of CodeBreaK 200 are robust, as the PFS hazard ratio withstands multiple sensitivity analyses. FDA agrees that the estimated PFS hazard ratio is generally consistent

across multiple analyses; however, we also note our tipping-point analysis, which showed that the statistical significance of the hazard ratio may not hold under different assumptions regarding the level of informative censoring caused by early dropouts and early crossover.

Both the applicant and the FDA agree that based on an interval censoring method, the PFS benefit could be as low as 5 [indiscernible] days. We note a higher rate of PFS events on the sotorasib arm, though we acknowledge this must be viewed in the setting of incomplete information with early dropouts on the docetaxel arm. When evaluating the Kaplan-Meier curves, we note that given high levels of censoring, the latter half of the curve, which appears to be separated, may not be reliable. This comprehensive assessment highlights uncertainty regarding the robustness of the PFS results and our ability to quantify the treatment effect of sotorasib.

In today's presentations, you will hear a discussion about both the design of CodeBreaK 200

and the conduct, and it is important to make the distinction between the two. Certain elements of the study design of CodeBreaK 200 -- in particular, the open label nature of the trial -- certainly may have influenced study conduct. Study conduct issues are those such as informative censoring and individual decisions, and patient management favoring sotorasib, which collectively represented a potential systemic bias impacting the fidelity of the primary PFS endpoint, as well as the overall trial results. You will hear from Dr. Vellanki later in the FDA presentation about several mitigation strategies, which may be utilized in open-label trials to help address expected bias.

When assessing whether the results of CodeBreaK 200 may be used to convert the accelerated approval of sotorasib to a traditional approval, we must consider whether the PFS per BICR results can be reliably interpreted. If so, then CodeBreaK 200 could potentially serve as confirmation of clinical benefit and fulfillment of the postmarketing requirement. We note for the

committee that a lack of superiority finding does not infer a noninferiority finding. This would require an a priori statistical design and assumptions, often involving relatively large patient numbers and most suited for an overall survival endpoint.

If CodeBreaK 200 cannot be used to verify clinical benefit, for example, if our concerns regarding study conduct supersede the narrow therapeutic effect of sotorasib relative to docetaxel, we would have an accelerated approval which has yet to be converted to a traditional or regular approval, and we would consider potential next steps within our regulatory framework.

FDA oncologists recognize the unmet need for patients with actionable mutation such as KRAS G12C, as well as evolving treatment paradigm. A decision to withdraw an accelerated approval is not automatic in the setting of a "failed" confirmatory trial; it is affected by many factors, all of which we will consider for sotorasib. We consider the nature of the "failed" trial. For example, if

there is a detriment in survival, we consider the current therapeutic landscape at the time of the failed trial, not at the time of the initial accelerated approval, and certainly we consider a potential safety advantage of the drug granted accelerated approval.

This is just a snapshot of what is publicly known regarding other drugs in development for patients with KRAS G12C mutations. The FDA has a wide-angle view of the therapeutic landscape, including other trials which may be ongoing or planned, and thus can reasonably assess areas of current or future unmet need. The furthest along in development is adagrasib, which was granted accelerated approval in December of '22, about 18 months after the approval of sotorasib. The sponsor is conducting a clinical trial very similar to CodeBreaK 200 called KRYSTAL-12. Some key differences from CodeBreaK 200 included a 2-to-1 randomization schema and crossover after real-time BICR was implemented from study start. This trial is ongoing, and the design certainly was influenced

by external trial results and anticipated open-label buying. FDA also notes that the applicant is planning another randomized trial in the first-line setting, CodeBreaK 202, which could potentially be used to confirm benefit.

As oncologists, we all seek to provide the best care for our patients, and in CodeBreaK 200, we believe that all parties were acting in what they perceived to be the best interest of patients based on available data; however, it is the collective pattern of conduct in this trial which raised concern about the fidelity of the primary endpoint, and in turn the overall trial results. We noted that our initial assessment of CodeBreaK 200 was equivocal; however, as signals of bias continued to emerge during the course of the review, we felt it was important to bring forth these issues to a wider oncology community.

The FDA is one of the only regulatory health authorities in the world who does patient-level data analysis from raw data sets, and thus would be in a position to perform high-quality objective

analyses of clinical trial data. We must go where the data takes us, sometimes even in spite of our own biases and enthusiasm for novel therapies. For patients in the U.S. and around the globe, through parallel regulatory reviews via OCE's Project Orbis, the FDA seeks to approve and label cancer therapeutics based on high-quality evidence which is robust and can withstand statistical pressure testing or various sensitivity analyses. And again, we note that while we often see bias in oncology trials, it may be able to be mitigated by factors such as endpoint selection and magnitude of benefit.

For the purposes of today's discussion, we ask the committee to discuss the multiple signals of potential bias favoring sotorasib, as well as patterns in study conduct in the context of top-line efficacy results. Did CodeBreaK 200 demonstrate superiority of sotorasib versus docetaxel? And if so, can we reliably quantify its effect? We will ask the committee to vote on whether the primary endpoint of progression-free

survival per blinded independent central review can be reliably interpreted in CodeBreaK 200.

Finally, I would like to thank the patients, caregivers, providers, research staff, and investigators involved in the study of sotorasib and other drugs in class. Trials like CodeBreaK 200 are designed to answer important scientific questions. We acknowledge the challenges such trials face and hope to spend more time discussing mitigation strategies moving forward. I look forward to a thoughtful discussion today between FDA, the applicant, and our advisory committee. Thank you.

DR. KLINE: Good morning. My name is Jackie Kline, and I'm the vice president of Global Regulatory Affairs for Oncology at Amgen. I'd like to thank the committee for their time today.

DR. MADAN: Dr. Kline, let me just introduce this portion here.

Both the FDA and the public believe in a transparent process for information gathering and decision making. To ensure such transparency at

the advisory committee meeting, the FDA believes that it is important to understand the context of an individual's presentation. For this reason, FDA encourages all participants, including the applicant's non-employee presenters, to advise the committee of any financial relationships that they may have with the applicant, such as consulting fees, travel expenses, honoraria, and interest in the applicant, including equity interests and those based upon the outcome of the meeting.

Likewise, FDA encourages you at the beginning of your presentation to advise the committee if you have any such financial relationships. If you choose not to address this issue of financial relationships at the beginning of your presentation, it will not preclude you from speaking.

We will now proceed with the Amgen, Incorporated presentation. Sorry for the interruption. Please continue.

Applicant Presentation - Jackie Kline

DR. KLINE: Thank you.

Good morning. My name is Jackie Kline, and I'm the vice president of Global Regulatory Affairs for Oncology at Amgen. I'd like to thank the committee for their time today, and the FDA for the opportunity to review important data for sotorasib from our phase 3 study, CodeBreaK 200.

FDA granted accelerated approval for sotorasib based on CodeBreaK 100, a single-arm study in patients with non-small cell lung cancer with the KRAS G12C mutation. Today, we will discuss the results of our confirmatory study known as CodeBreaK 200, which provides a head-to-head comparison of sotorasib to docetaxel. FDA has raised important questions about the reliability of the results of this study. Amgen will present data and analyses that address FDA questions and demonstrate that CodeBreaK 200 can be reliably interpreted to confirm the clinical benefit of sotorasib.

Lung cancer is the leading cause of cancer-related deaths worldwide. In the United States, the KRAS G12C mutation is estimated to be

present in approximately 13 percent of patients with lung adenocarcinoma. This equates to an estimated 10,000 patients with advanced disease. This mutation impairs cycling of KRAS and leads to oncogenic signaling and subsequent tumorigenesis. Notably, it rarely occurs in the presence of other actionable mutations.

Sotorasib is a first-in-class therapy that covalently binds to the KRAS G12C mutated protein, locks it in the inactive state, and prevents downstream proliferation and signaling. CodeBreaK 200 is a phase 3 study designed to confirm the clinical benefit of sotorasib. Initially, the study was designed with a primary endpoint of progression-free survival and several key secondary endpoints, including overall survival. To provide sufficient power for assessment of overall survival, the study was designed to enroll 650 patients.

While CodeBreaK 200 was enrolling, the primary analysis for CodeBreaK 100 was completed. In that study, sotorasib demonstrated an objective

response rate of 36 percent and a median duration of response of 10 months. Based on these results and in consultation with FDA, Amgen decided to focus only on the progression-free survival primary endpoint and amended the study protocol to decrease the sample size to 330 patients. With this change, the power for the overall survival endpoint was substantially decreased. In the same amendment, patients were eligible for crossover from docetaxel to sotorasib upon documentation of progressive disease. While the enrollment of patients was close to completion when this amendment was finalized, only 25 percent of progression events had occurred by that time.

Today, we will present data to support the following points. Treatment with sotorasib results in improved progression-free survival over docetaxel and rapid and durable tumor response. Sotorasib exhibits a differentiated safety profile as compared to docetaxel. Risks are well characterized and manageable. CodeBreaK 200 can be reliably interpreted to confirm the clinical

benefit of sotorasib. And finally, sotorasib provides an important option for the treatment of patients with KRAS G12C mutated non-small cell lung cancer.

In addition to our presenters, we also have Dr. Gary Koch and several Amgen subject matter experts available to answer questions. And now, Dr. Mehta will present the efficacy results from CodeBreaK 200.

Applicant Presentation - Bhakti Mehta

DR. MEHTA: Thank you, Dr. Kline.

Good morning. My name is Bhakti Mehta. I'm an executive director within the oncology clinical development group at Amgen. Today, I will review the efficacy data from CodeBreaK 200.

CodeBreaK 200 is a global, randomized phase 3 trial of sotorasib versus docetaxel in patients with non-small cell lung cancer. Key eligibility criteria included KRAS G12C; locally advanced and unresectable or metastatic disease; at least one prior systemic therapy for advanced disease, including platinum-based chemotherapy and

checkpoint inhibitors, given either as one line of therapy or as separate lines.

While patients with active brain metastases were excluded, patients with previously treated brain metastases were eligible. Patients were stratified based on prior lines of therapy, race, and history of CNS involvement. Patients were randomized 1 to 1 to either sotorasib, given as 960 milligrams oral once daily, or docetaxel, given as 75 milligrams per meter squared intravenously every 3 weeks. Response assessment scans were performed every 6 weeks on both arms.

While physicians and patients knew their assigned treatments, the sponsor and the imaging vendor were blinded to treatment assignment until the primary analysis. The primary endpoint was progression-free survival as assessed by blinded independent central review. Secondary endpoints included overall survival; response rates; duration of response; time to response; disease control rate; patient-reported outcomes; and safety assessment. The protocol was amended to allow

patients on the docetaxel arm who had centrally confirmed radiological progression to crossover to sotorasib.

From June 2020 to April 2021, 345 patients from 148 centers in 22 countries were randomized to receive either sotorasib or docetaxel. Two patients randomized to sotorasib and 23 patients, or 13 percent, randomized to docetaxel withdrew before receiving study treatment. Dr. Friberg will address this imbalance in early dropouts later in the presentation.

The most common reason for treatment discontinuation was disease progression on both treatment arms. At the time of the data cutoff, 22, or 13 percent, of sotorasib patients, and seven, or 4 percent, of docetaxel patients, were still receiving the assigned treatment. Approximately a quarter of patients randomized to docetaxel crossed over to sotorasib on protocol, with the further 7 percent of patients known to have received sotorasib off study.

The baseline characteristics were generally

well matched between the arms and reflective of the KRAS G12C patient population. The median age for both groups was 64. Nearly all were smokers and were predominantly ECOG performance status 1. One-third of patients had a history of CNS involvement, and approximately a fifth of patients had liver metastases. The median prior lines of therapy were 2, and per protocol, patients were required to have received both platinum-based chemotherapy and checkpoint inhibitors.

The study met its primary endpoint of progression-free survival by blinded independent central review at a median study follow-up time of 17.7 months. With sotorasib in blue and docetaxel in gray, you see an early separation of the curves, starting at the first scan and sustained throughout the course of the follow-up.

The most informative method of assessing the PFS benefit is to look at the hazard ratio, which looks at the entirety of the Kaplan-Meier curve. As represented by the green area, sotorasib demonstrated superiority over docetaxel, with a

statistically significant hazard ratio of 0.66 and a p-value of 0.003. This represents an estimated 34 percent average lower risk of an event of progression or death with sotorasib compared to docetaxel.

Medium PFS values are one measure of the differences between the treatment arms, albeit [indiscernible] only one point on the Kaplan-Meier curve, the 50th percentile on the Y axis. The median was 5.6 months in the sotorasib arm versus 4.5 months in the docetaxel arm. At the 1-year milestone, the PFS rate for sotorasib was 25 percent versus 10 percent for docetaxel. This measure represents the effect size more robustly, as this vertical difference between the two curves is similar across the ranges from 8 to 14 months.

Now, how do these PFS results look in different prespecified subgroups? Here is a forest plot of the PFS hazard ratios that shows all the point estimates to the left, indicating that the hazard ratios remain in favor of sotorasib over docetaxel across subgroups, including demographics,

performance score, prior lines of therapy, and in poor prognostic groups such as history of CNS involvement and liver metastases.

We will now examine several sensitivity analyses that were conducted to test the robustness of the data. Included in these analyses are investigator-assessed PFS and three additional analyses recommended by regulatory guidelines: first, the investigator-assessed PFS, which does not suffer from censoring or other issues that may be attributed to central imaging; second, an analysis in which all patients who initiated new anti-cancer therapies were considered progressed at the time of the anti-cancer treatment switch; third, an analysis treating any withdrawal of consent or loss of follow-up as a PFS event; and fourth, an analysis using the scheduled scan assessment date instead of the actual assessment date.

As we can see in the right column, all of these prespecified sensitivity analyses resulted in hazard ratios consistent with the primary analysis

and favored sotorasib over docetaxel. These are described in greater detail in the briefing document. Taken together, these analyses show that the primary endpoint outcome of PFS was robust.

Turning to additional efficacy endpoints, there was an approximately 2-fold higher centrally confirmed response rate for sotorasib versus docetaxel. ORRs were 28 percent and 13 percent, respectively, which met the predefined threshold for statistical significance, with a p-value of less than 0.001. This waterfall plot presents the best confirmed change in target lesion size for sotorasib on the left and docetaxel on the right. The disease control rate was also higher for those on sotorasib versus docetaxel, 83 percent versus 60 percent.

Now, let's look at the time course of these responders in greater detail. There were 48 responders on the sotorasib arm, 2-fold more than the 23 responders on the docetaxel arm. A patient's first response is indicated by an orange circle, a red cross mark indicates data progression

or death, and a green arrow indicates an ongoing response. There was a 6-week faster median time to response for sotorasib compared to docetaxel, 1.4 months versus 2.8 months. The median duration of response was longer with sotorasib, 8.6 months versus 6.8 months. These data show that sotorasib treatment led to twice as many responders, with the responses occurring in half the time and lasting 2 months longer.

Turning now to overall survival, which was similar between the treatment arms, the Kaplan-Meier curves are overlapping, with a hazard ratio of 0.96 in this updated data cut. Now, let us examine the OS results in different subgroups. This forest plot for overall survival shows all the hazard ratio point estimates right down the middle, indicating that the OS was similar in both arms across all subgroups.

Now, we turn to patient-reported outcomes. Patient-reported outcomes were not formally statistically tested due to hierarchical testing rules; however, we believe that it is important to

share the patient experience with these treatments. In CodeBreaK 200, patients were asked to complete these well-established PRO questionnaires to capture the perception of quality of life and symptom burden. PROs were measured at baseline and on day 1 of each subsequent cycle until treatment discontinuation. The analyses' endpoints were change from baseline to week 12, time to deterioration, and descriptive statistics.

I will review the data on the time to deterioration PRO measures. Dr. Eisele will present data on the patient experience with side effects from the FACT-G measure, and a comprehensive review of the other PRO measures are provided in the briefing document.

Here, you see the median time to deterioration in weeks in the quality-of-life measures of global health status and physical functioning and the time to deterioration for the symptoms of dyspnea, cough, and chest pain. In all of these PRO measures, sotorasib delayed the time to deterioration compared to docetaxel.

To conclude a review of the efficacy data, sotorasib showed significant improvement in the primary endpoint of progression-free survival versus docetaxel. PFS benefit was consistent and statistically robust between central and investigator review across subgroups and in prespecified sensitivity analysis. The overall response rate, the disease control rate, time to response, and duration of response were all improved for sotorasib versus docetaxel. Overall survival was similar. Patient-reported outcomes favored sotorasib across a variety of measures.

I'll now hand it over to Dr. Eisele to review the safety data. Thank you.

Applicant Presentation - Osa Eisele

DR. EISELE: Thank you, Dr. Mehta.

Good morning. I'm Osa Eisele, executive medical director within Global Patient Safety, and I'll be reviewing the CodeBreaK 200 safety data.

The safety profile of sotorasib is supported by a robust data set that includes over 2,000 patients who have received sotorasib in the

clinical development program and over 5,000 patient-years of postmarketing exposure. The monotherapy safety data shown here, and which includes the CodeBreaK 200 study, is a subset of the clinical development program. The CodeBreaK 200 safety data set consists of 169 sotorasib patients and 151 docetaxel patients.

The median duration of treatment in the sotorasib arm was longer, at 20 weeks and 7 cycles, compared to 12 weeks and 4 cycles in the docetaxel arm. The median relative dose intensities of sotorasib and docetaxel were comparable and consistent with their targeted doses. The longer duration of study treatment for sotorasib with similar relative dose intensities across both treatment arms speaks to the overall tolerability of sotorasib.

In this study, nearly every patient experienced at least one adverse event. Grade 3 or higher adverse events were more frequent with sotorasib, while grade 4 or higher adverse events were more frequent with docetaxel. The incidences

of fatal adverse events, serious adverse events, and adverse events leading to discontinuation were similar between treatment arms. Events leading to dose modification were higher with sotorasib, and this was driven by more frequent dose interruptions.

This double tornado plot shows adverse events with at least a 5 percent difference between the two treatment arms. The plot on the top reflects side effects that are more common with sotorasib, while the bottom reflects those more common with docetaxel. Diarrhea and elevations in liver tests were more frequent with sotorasib, and this is consistent with its safety profile. For docetaxel, common events, including fatigue and alopecia, are also consistent with its safety profile. For sotorasib, the most frequent grade 3 or higher adverse events were diarrhea, ALT, and AST elevations. For docetaxel, these include fatigue, pneumonia, and neutropenia. These events are again consisted with each drug's established safety profile.

Treatment modifications, which include dose interruptions and reductions, are the main strategies to effectively manage toxicities. Importantly, amongst sotorasib patients with dose interruptions due to adverse events, the total treatment interruption duration was less than 10 percent of the total treatment duration. As seen in the table, the types of events leading to treatment modification were, again, consistent with each drug's safety profile.

Turning now to serious adverse events, hospitalizations in advanced cancer patients are common, and in this study, almost half the patients in either treatment arm were hospitalized, mostly due to their underlying cancer. For sotorasib, the most frequent events leading to hospitalization were hepatotoxicity and diarrhea, though collectively, these accounted for a small proportion of all hospitalizations. For docetaxel, the most frequent events were infections, neutropenia, and anemia, and these accounted for a much larger proportion of all hospitalizations.

The incidence of treatment-related hospitalizations was also high in docetaxel, at 22 percent versus 9 percent for sotorasib. In summary, docetaxel toxicities more frequently resulted in hospitalizations than sotorasib toxicities.

Now, we'll move on to discuss the sotorasib key risks. These are risks that are important from a clinical perspective in terms of either frequency or severity, and include diarrhea, hepatotoxicity, and interstitial lung disease. While ILD can be severe, it is infrequent and is described in detail in the briefing book. Thus, the remainder of the presentation will focus on diarrhea and hepatotoxicity.

Diarrhea occurred in 41 percent of patients, and the majority of events were grade 1 and 2, shown here in green and blue. There were no grade 4 or fatal diarrheas. The table to the right shows management and outcomes of diarrhea. Diarrhea was effectively managed with dose interruptions in 15 percent of patients and dose reductions in 8 percent, with only one patient

discontinuing treatment. Use of antidiarrheals was reported in 76 percent of these patients. Events fully result in the majority of patients and the median duration of diarrhea was 22 days. In summary, the data demonstrate that diarrhea is tolerable and manageable.

Now, we will review the hepatic events, which occurred in 24 percent of sotorasib patients. As evidenced by the most commonly reported AEs in the table on the left, these events are primarily characterized by ALT/AST elevations. Most of the events were grade 3 severity, shown in orange on the right. Importantly, there were no cases of severe liver injury with hepatic failure and no fatal events.

Shown here is the time course of ALT and AST for each patient whose transaminase levels were greater than 3 times the upper limit of normal. Blood chemistry was collected on day 1 of each cycle. As we can see by the peaks, in the majority of patients, ALT and AST elevations were below 10 times the upper limit of normal. The declines

from peak elevations coincide with treatment modification and speaks to both the reversibility and manageability of these lab abnormalities. Management of hepatic AEs was primarily through treatment interruption in 18 percent of patients and dose reductions in 7 percent of patients. In 8 percent of patients, treatment was withdrawn.

Approximately 70 percent of these patients were also administered steroids. For the majority of patients, events fully resolved and the median duration of these events was 22 days. In summary, the data demonstrate that hepatic events can be managed through those modifications and supportive care.

Now switching to patient-reported outcomes and specifically the FACT-G item GP5 named, "I am bothered by the side effects of treatment." This PRO is a validated tool and is a summary measure of side effect impact to the individual subject. Now, let's look at the stacked bar chart for response rates for sotorasib on the left and docetaxel on the right. The yellow, orange, and red bars

illustrate patients who were more bothered by their side effects, and these are more prevalent in the docetaxel chart compared to the sotorasib chart.

In conclusion, the safety data from CodeBreaK 200 was consistent with the known safety profile of sotorasib. The safety profile of sotorasib and docetaxel are characterized by different types of adverse events, and sotorasib patients report being less frequently bothered by their side effects. Lastly, key risks of sotorasib can be effectively monitored and are manageable with treatment modifications and supportive care.

Thank you. I'll now hand the presentation over to Dr. Friberg.

Applicant Presentation - Gregory Friberg

DR. FRIBERG: Thank you, Dr. Eisele, and good morning. My name is Greg Friberg. I am a vice president of Medical Affairs at Amgen. The major question before you today is whether CodeBreaK 200 can be reliably interpreted. You're being asked to judge whether these results are believable and whether they can be trusted, given

potential sources of bias.

To address this directly, let's review the concerns highlighted in the FDA briefing document. They list criteria A through F, which define what is needed for a study to be considered adequate and well controlled. Similarly, they summarize four overarching areas of concern, which call into question whether CodeBreaK 200 meets these criteria.

First, there were high rates of early dropout on docetaxel. The concern here is that the effects of randomization were lost and that the arms are no longer comparable. Second, there were discrepancies between investigator and central reads for progression. The implication here is that this was the symptom of a larger problem; that investigator choices caused premature censoring of docetaxel patients. Third, quality measures relating to the central reads call into question the reliability of these assessments altogether. Finally, there is a concern that all of these issues, when compounded, challenge whether the

primary endpoint can be believed.

These four areas of concern raise fair and appropriate questions. They need to be thoroughly interrogated to ensure that we have confidence in the CodeBreaK 200 results. We will do so one at a time and present several additional analyses to provide some context.

First, we will focus on early dropouts. We will walk through several approaches which attempt to restore confidence in the randomization. Next, we will discuss imaging procedures, and we'll review the steps that were taken to ensure reliability. Third, we will dive into the potential impact of read discrepancies, specifically as related to event censoring.

This was an open-label study, and it is possible that investigators behave differently based upon their knowledge of treatment. We will explore this potential impact with several tipping-point analyses using different levels of statistical pessimism. Finally, we will focus on a more holistic question. Do the CodeBreaK 200

results clinically make sense in the context of other studies?

With respect to the early dropout, we know that 23 patients randomized to docetaxel withdrew from the study without ever receiving treatment. When asking how this may have affected the results, it is essential to know who these patients were. We know they were recruited from 21 different sites and 13 different countries. They did not appear to be influenced by individual investigators.

Clinical covariates for these 23 patients are shown in the table. If anything, these patients had less favorable profiles than those who actually received docetaxel. The factors highlighted in green show that they had worse performance status and a higher percentage of brain or liver metastases. These early dropouts did not clinically appear to be those destined for the best outcomes.

As a sensitivity analysis, we performed a stratified Cox model adjusted for clinically relevant covariates to address the imbalances in

baseline characteristics introduced by early dropouts. Five covariates were selected using clinical considerations and a required prevalence of at least 10 percent. As a methodologic reminder, this approach already accounts for the prespecified stratification factors on the right.

The PFS hazard ratio after this adjustment was 0.60, with confidence intervals shown favoring sotorasib. The fact that this hazard ratio was slightly more favorable for sotorasib as compared to the primary analysis supports the clinical observation from the baseline characteristics that early dropout patients may have had a less favorable prognosis.

While the Cox model is reassuring, we wanted to lean on these results further to better understand the impact from early dropouts. You have seen a version of this figure before in the efficacy section. The 23 untreated early dropouts on the docetaxel arm are circled in red. We wanted to simulate how these patients might have performed had they not dropped out. To do this, we sampled

from the pool of patients circled in green at the bottom of the figure. These 120 patients are from the same randomization pool and have remained on study for at least 6 weeks. Sampling from this patient pool accounts for both measured and unmeasured variables.

This also removed some of the poorest prognosis patients who were unable to reach week 6 due to progression, death, or censoring. We performed multiple trial simulations, randomly sampling in strata patients to replace each of the 23 untreated early dropouts. We chose not to impute new results for the two early sotorasib dropouts.

Here you see the results of this imputation exercise. The average PFS hazard ratio across 20,000 simulations was 0.70, with the overwhelming majority, over 99 percent, demonstrating PFS superiority for sotorasib. While we acknowledge the imbalance caused by untreated docetaxel dropouts, these sensitivity analyses are reassuringly consistent with the primary result.

While these imputations use actual study data, other imputation methods can instead choose varying levels of optimism or pessimism and ask how extreme would one's assumptions need to be in order to change the overall result? In table 12 of the FDA briefing document, such an approach was taken. Early dropouts were replaced with patients sampled from the top 50 percent of PFS times. It is fair to ask what this actually means in practical terms and how extreme is this assumption?

To add some perspective, this figure presents the real PFS curve for docetaxel patients on the CodeBreaK 200 study, and here is the curve for the docetaxel patients with the 50 percent best PFS times. This curve is generated using the real data. As you can see from their table 12, the FDA modeling produces an even longer median PFS for docetaxel. Imputation for early docetaxel dropouts, using this more optimistic pool of patients, was consistent with the CodeBreaK 200 primary analysis. The hazard ratio was 0.73 and the confidence intervals excluded 1.

Now, here is the curve for the top 50 percent of PFS times from both arms of the study pooled together. Again, the medians for the FDA modeling are even longer. After imputing early dropouts with patients sampled from this most optimistic pool, the results again favor the experimental arm. The hazard ratio was 0.77, but now the confidence interval just tips past 1. It is fair to ask how realistic the docetaxel assumptions are in these analysis. These are a very optimistic pool of patients. This is an extreme assumption about how the dropout patients might have performed, and the fact that the sotorasib hazard ratio remains superior is actually quite reassuring.

Now, let's focus on the imaging reads and how the primary endpoint came to be based upon a 100 percent BICR re-read. Here is what occurred. During the execution of the study, sponsors routinely performed aggregate event projections in order to determine the timing of the primary analysis. In early 2022, the BICR-determined

timing projections for CodeBreaK 200 became unstable. Amgen identified that a discordance was present between aggregate confirmation of progression, or COP event numbers, and aggregate BICR event numbers.

This discordance, without any knowledge of treatment assignment, was communicated to the imaging vendor. The imaging vendor initiated the quality review in accordance with the charter. The results of this independent review resulted in the imaging vendor retraining their radiologist, and the reads for 11 subjects were corrected. Amgen discussed this series of events with the FDA. The FDA recommended, and Amgen agreed, to perform a 100 percent re-read of all primary images to nullify any potential bias.

This was a conservative step. It was both appropriate and thorough. All procedures adhered to the imaging charter and the suggestion of violations are inaccurate. Given this 100 percent re-read with a new team of radiologists, it is difficult to imagine how these events could have

influenced the primary analysis.

Let's shift gears and focus on the possible impact of investigator actions on censoring events. The FDA has rightly noted that investigators may have called early progressions on the docetaxel arm perhaps in their enthusiasm to switch patients to sotorasib. This figure accounts for all the patients on CodeBreaK 200. You see the total number of BICR PFS events at the top; below you see the censoring in each arm. There were a total of 49 censoring events for sotorasib and 73 for docetaxel.

Now, we've already discussed the early dropouts, so I want to focus on the green box, which counts patients censored for the start of new anti-cancer therapy. These patients were started on new treatments by their doctors, but the independent reader did not agree with the local PD call. This group includes the 19 early crossover patients noted in the agency's briefing document.

The next set of analyses will address the potential bias introduced by these 24 and

31 censored patients. The question again is whether premature censoring when no BICR PFS event had occurred influenced the primary PFS analysis. Sensitivity analyses here allow us to ask, how extreme would our assumptions need to be to render the primary outcome unreliable?

Let me walk you through two tipping-point analyses. Here, we assume the worst for sotorasib, that all 24 patients experienced the PFS event on the day of the new therapy. We pessimistically count them all as progressors, then we assume the best for the docetaxel patients and work in the other direction. We optimistically assume that none of these 31 patients were real progressions and ask, how many censored patients would need to be real in order to restore our confidence in the overall finding?

The results of this exercise are visualized as dots for the hazard ratio and whiskers for the confidence intervals. Moving left to right, the imputation exercise adds docetaxel progressions one by one, starting with zero and ending with all 31.

In every scenario, you can see that the hazard ratio dot favors sotorasib and only the most extreme scenario does the whisker cross 1. Even when all 24 sotorasib patients are considered progressors, if just one docetaxel patient experienced a reliable PFS event, as was called by their treating physician, then the confidence interval excludes 1. This analysis shows us that you would have to make extreme assumptions to render the result no longer significant.

Now, how about if we compound the censoring issues? How would the results be affected if we used this approach to account for both therapy switching and untreated early dropouts? Again, we pessimistically assume that all 26 censored sotorasib patients were progressors and we optimistically assume that none of the 51 censored docetaxel patients were.

Here, we make extreme, arguably, unrealistic assumptions, and yet the hazard ratio dot still favors sotorasib in all scenarios. When all 26 patients are considered progressors, adding back

just 3 events from the 51 censored docetaxel patients tips the confidence interval to exclude 1. By definition, these analyses are intended to explore extreme scenarios, and indeed it is only in such extreme settings that the benefit for sotorasib is called into question.

Now, the ultimate question today is whether you can trust the CodeBreaK 200 results as real? One additional way to address this is to ask how the CodeBreaK 200 data compares to other clinical trials, especially around the PFS primary endpoint. In spite of the potential challenges discussed, this trial reports clinical results that are remarkably consistent with other trials for both sotorasib and for docetaxel. These results were generated on different trials, in different regions, and at different time points. Furthermore, these data are consistent with real-world evidence for G12C-specific populations.

This consistency gives further assurance that in spite of the aforementioned challenges, the results of the study are indeed interpretable.

Sotorasib delivered an improved PFS when directly compared to docetaxel. To provide further insights from a physician and patient perspective, I would like to hand the podium over to Dr. Melissa Johnson.

Applicant Presentation - Melissa Johnson

DR. JOHNSON: Thank you, Dr. Friberg.

My name is Melissa Johnson, and I'm the director of the Lung Cancer Research program at Sarah Cannon Research Institute in Nashville. I'll be discussing my clinical perspective on the information you've seen today. My institution has been compensated for my time and I have no financial interest in the outcome of this meeting.

There's no question in my mind that immunotherapy has transformed my oncology practice and heightened the expectations of my lung cancer patients. One out of five non-small cell lung cancer patients experience long-term overall survival benefit. That means the majority of patients who come to see me are hoping for and now need more from their treatment.

While these immunotherapies have advanced care, here's what shouldn't be overlooked about these advances. Looking at the KEYNOTE-189 progression-free survival curve, over 60 percent will progress by one year with our current standards, and we'll need a second-line treatment. Moreover, with immunotherapy being used in the first-line setting, clinical benefit with it in subsequent lines of therapy is lacking. We have to do more for patients.

The majority of immunotherapy refractory lung cancer patients who actually receive second-line treatment will be treated with docetaxel. Historically, chemotherapy options beyond this are limited and have minimal efficacy. Importantly for our topic today, in patients with KRAS G12C mutations, these remain the only options outside of drugs like sotorasib or clinical trial, and here's the truth about docetaxel. It's an active drug, and that's why it continues to be our globally recognized second-line standard and clinical trial comparator with median

progression-free survival of 3-to-5 months, median overall survival, 8-to-10 months, and an objective response rate of 10 to 15 percent. However, that benefit doesn't come easily. It's neither an easy drug to give, nor is it always tolerated. In fact, clinicians love to hate this drug, and patients dread it, which is why I'm not surprised to see the early dropout in CodeBreaK 200.

Docetaxel is dosed intravenously once every 3 weeks. It's frequently dose-reduced for tolerance. It requires 3 days of oral premedications, 8 milligrams of dexamethasone -- quite a large dose -- twice daily to reduce the risk of hypersensitivity and fluid retention complications. In my patients, it causes distressing side effects: febrile neutropenia, stomatitis with impaired oral intake, alopecia and nail changes; nausea, vomiting and asthenia.

From the CodeBreaK 200 trial, here are four key findings, all impactful, that resonate with me. Said another way, these data give me clinical confidence to use sotorasib. First, the

Kaplan-Meier curve estimating PFS. If anything, the patients receiving docetaxel did better than expected, which we believe may be linked to prior IO use. The small improvement in the medians has drawn focus; however, medians are only one way to measure clinical benefit. Here, we see that the curves separate at the first scan and stay separated for the duration of the trial. To me, that means that at every point along this curve, we can measure benefit for patients receiving sotorasib.

Next, who are the patients who drove this benefit? All of them. You see that nicely from the forest plot, that across all subgroups, the blue dots are shifted to the left. To me, this slide illustrates the fact that sotorasib is a targeted inhibitor, selective for all patients with KRAS G12C mutations and more capable of controlling the rate of disease growth than the non-selective chemotherapy docetaxel.

Also, more tumor response with sotorasib, 83 percent versus 60 percent achieving disease

control. This is the endpoint that patients feel and understand. It equates to sotorasib's higher chance of abating cancer-related symptoms. In fact, these endpoints dovetail nicely with patient-reported outcomes. Patients are less bothered by their side effects. These are also important measures and align with what my patients tell me they are experiencing.

While these are all objective trial endpoints, taken together, there is a very real subjective difference in what patients are experiencing with sotorasib versus docetaxel. They feel it, and I can see it. The improvements that we intuit as physicians, managing patients with targeted therapy versus chemotherapy cannot be overstated.

Safety is sometimes harder to meaningfully illustrate. What might not be obvious from this tornado plot is just how different these drugs are to a practicing oncologist, let alone a patient. Managing docetaxel-related side effects is complicated. Patients go home with an on-body

injector or they come back to clinic for Neulasta to protect against febrile neutropenia and hospitalizations. I bring patients back to the clinic the second week, and sometimes the third week of the cycle, to manage dehydration that comes with nausea and stomatitis. I send patients for transfusions for anemia. I write prescriptions for wigs.

Navigating sotorasib side effects is easier. Patients come back to clinic periodically for liver function tests. If they're feeling well, they don't come back until their first scan at 6 weeks. Patients go home with a script for Imodium for diarrhea and Zofran for nausea, in addition to sotorasib, and often manage their symptoms themselves. They take their pills at home, they go back to work, and they enjoy more independence from the clinic, and no one has to know that they have cancer because they don't lose their hair. It should be their choice who knows about their illness.

Let me leave you with this. Patients prefer

oral medications. Once daily dosing is an added plus. Sotorasib has been criticized for not beating docetaxel on overall survival and it's true that the Kaplan-Meier curves are very similar. My patients want to live longer, but if they can live the same amount of time and live better, as all the PFS safety and PRO endpoints demonstrate, they will pick sotorasib every time, and so will their doctors who are helping them make these decisions.

Patients need options beyond docetaxel. Patients should absolutely be able to choose a well-tolerated oral therapy designed to inhibit their driver oncogene in lieu of an unselective IV chemotherapy and its liabilities. I've used sotorasib as a well welcomed addition to my armamentarium for the treatment of KRAS G12C mutated non-small cell lung cancer. I believe it is a step forward towards offering our patients more; more treatment options, more quality in their lives, and more control of their cancer's growth.

I will now turn it back over to Dr. Friberg. Thank you.

DR. FRIBERG: Thank you, Dr. Johnson, and thank you for your attention. We will look forward to answering your questions.

DR. MADAN: Thank you for that presentation from the sponsor. We will now proceed with the FDA's presentation, starting with Dr. Jeevan Puthiamadathil.

FDA Presentation - Jeevan Puthiamadathil

DR. PUTHIAMADATHIL: Good morning. I'm Dr. Jeevan Puthiamadathil, medical oncologist on the thoracic and head and neck cancers' team at the FDA. This presentation reflects the collective input of our FDA review team.

Dr. Singh in her opening remarks discussed FDA's rationale for convening today's advisory committee meeting. The FDA review team has found it challenging to interpret the results of CodeBreaK 200. The FDA believes that patient and investigator awareness surrounding the development and early response rates of sotorasib for patients with KRAS G12C mutated non-small cell lung cancer may have led to patterns in the study conduct

indicative of potential bias in favor of sotorasib.

Bias is not uncommon in randomized clinical trials or unique to CodeBreaK 200; however, in light of an incremental progression-free survival effect of 5 weeks and no difference in survival relative to a marginal comparator, these patterns of bias have led to uncertainty in our ability to interpret the primary PFS endpoint.

As part of our review framework, FDA aims to determine whether a trial is adequate and well controlled as defined by Title 21 of the Code of Federal Regulations. CodeBreaK 200 may lack certain features of an adequate, well-controlled trial, including adequate measures to minimize bias in subject assignment to treatment group to assure comparability of the groups; adequate measures to minimize bias in the parts of subjects, observers, and analysts of data; well-defined and reliable methods to assess response; and ultimately, adequate analysis of the results to assess the effect of the drug.

Our FDA review of CodeBreaK 200 suggested a

potential pattern of systemic bias and study conduct issues. While the trial was being conducted, the applicant triggered a review by the imaging vendor, which resulted in radiologic re-reads of patient scans, changing the PFS interim analysis from statistically not significant to statistically significant. FDA views the applicant triggering this process as a potential interference in imaging assessments and a potential violation of the imaging charter.

Later, our initial review of top-line results identified 23 patients randomized to the docetaxel arm who never received treatment, compared to only two on the sotorasib arm. Most of these patients did not receive study therapy due to patient request or withdrawal of consent. This asymmetric dropout led to the potential loss of the benefits of randomization.

Finally, during our review, FDA identified evidence of investigator assessments of imaging consistently favoring the sotorasib arm. These multiple signals of potential bias, systemic bias,

may have impacted our ability to adequately analyze the study results, which is a key feature of an adequate and well-controlled trial. We will ask the committee to discuss and vote whether the primary endpoint of PFS per BICR can be reliably interpreted.

Well before its accelerated approval of May 2021, early press for sotorasib fueled public awareness of the drug, touted as a breakthrough for patients with KRAS G12C mutated cancers who had long awaited the promise of precision medicine. As early as June 2019, the first clinical data was announced at the American Society of Clinical Oncology annual conference, about a year prior to the first patient being enrolled on CodeBreaK 200.

During the conduct of CodeBreaK 200, the public became aware of positive top-line results of sotorasib, as well as its breakthrough therapy designation. These public milestones could have led to a perceived loss of equipoise in CodeBreaK 200, with patients and investigators alike trying to gain access to sotorasib.

To enable a discussion about equipoise in randomized trials, here we provide several definitions. Equipoise is defined as the absence of certainty about which intervention is better. It is considered necessary for the ethical conduct of a randomized trial. Loss of equipoise occurs when there is certainty that one intervention is better than the other.

For this discussion, we consider the perceived loss of equipoise as the belief that one intervention is better, even without definitive evidence. When there is a perceived loss of equipoise, behaviors of trial participants, including patients and investigators, can change. In CodeBreaK 200, given today's information age likely resulting in widespread public awareness of sotorasib, even before the trial started enrolling, there may have been such perceived loss of equipoise.

The results of CodeBreaK 100, the single-arm trial evaluating sotorasib, eventually led to an FDA accelerated approval in May 2021, based on an

objective response rate of 36 percent with substantial durability. For a drug to be granted accelerated approval, there should be substantial evidence of effectiveness, the endpoints should be reasonably likely to predict clinical benefit, and there should be a therapeutic benefit over available therapy.

Given the historically low response rates of docetaxel in the second-line treatment setting, sotorasib clearly fell into the paradigm of an accelerated approval. The applicant proposed that CodeBreaK 200 served as a confirmatory trial to verify benefit of sotorasib in a randomized setting versus docetaxel. Given that KRAS G12C is the most common actionable oncogenic alteration identified in lung cancer and randomized trials are feasible and appropriate, FDA supported this development strategy.

CodeBreaK 200 utilized an open-label design. Patients were randomized to either single-agent oral sotorasib given daily or intravenous docetaxel given every 3 weeks. The primary endpoint of

CodeBreaK 200 was PFS per BICR. Crossover was instituted late in the trial after 99 percent of patients had been enrolled.

PFS has been commonly used to support approvals in oncology, particularly for targeted therapies; however, the PFS endpoint is inherently subject to some degree of bias. The criteria for disease progression are based on subjective interpretation of radiographic images in clinical evaluation. As such, there are several uncertainties in measuring PFS, including variability and timing of assessments and intra- and inter-reader variability.

As a result, PFS assessments are subjective interpretations with potential to introduce bias, particularly when used in open-label trials. This is in contrast to overall survival, which is a more objective endpoint and often considered the gold standard in oncology trials. For any trial with a primary PFS endpoint, FDA conducts sensitivity analyses to explore the strength of the primary analysis. The robustness of the treatment effect

should be seen across various measures of the endpoint, including the hazard ratio, medians, shape of the Kaplan-Meier curves, and event rates.

The median progression-free survival benefit of sotorasib was 5 weeks. This was statistically significant but small in magnitude, and less than the imaging interval of 6 weeks, raising concerns that the result could be lower using interval censoring. We also note more PFS events on the sotorasib arm compared to docetaxel. There was no difference in overall survival, and at the time of the primary analysis, 26 percent of patients from the docetaxel arm had crossed over to the sotorasib arm. The difference in objective response rate was statistically significant.

Patient disposition showed a high differential on patients who were randomized and not treated in the docetaxel, 23, relative to the sotorasib arm, two. Most of these patients from the docetaxel arm were not treated due to patient request or withdrawal of consent. It is conceivable that patients randomized to docetaxel

would either decide to receive docetaxel off trial with their local oncology provider or seek access to an alternative KRAS G12C inhibitor through another trial. This asymmetric dropout suggests the potential for investigator and patient bias favoring sotorasib. This pattern of behavior led to a loss of information and may have led to informative censoring of PFS results.

In randomized-controlled trials, blinding helps minimize bias by preventing patients and study personnel from gaining knowledge of treatment arm assignment. Blinding is feasible in certain therapeutic settings; however, in oncology, an open-label design is often necessary because of differences between trial arm interventions, such as route of administration and side effect profiles.

This was the case for CodeBreaK 200. An open-label design is susceptible to bias, particularly when the standard-of-care treatment used in the control arm is thought to be suboptimal. Docetaxel has a historic response rate

of 8 to 12 percent versus a 36 percent response rate seen in the single-arm trial of sotorasib. Systemic biases are difficult to prove, but data may signal their presence. In CodeBreaK 200, we identified asymmetric early dropout and investigator imaging assessments favoring the sotorasib arm as signals for potential systemic bias. It is unknown what data were not captured due to potential underreporting of adverse events and patient-reported outcomes, both of which are subjective data elements.

We will now discuss a review of the efficacy and safety results. The Kaplan-Meier curves for PFS show an initial separation, suggesting sotorasib may have a treatment effect over docetaxel; however, this initial separation decreases as the curves come together at about 7 months. While it is noted that there is a greater separation after 7 months, which may indicate a greater benefit for sotorasib, the curves start to come back together again around 15 months, potentially negating long-term

superiority of sotorasib over docetaxel.

Additionally, after about 7 months, there are relatively few patients remaining in follow-up who have not been censored, as shown in the red box; therefore, the separation in these curves cannot be reliably interpreted. This, along with the median follow-up of only 6.9 months in the docetaxel arm, creates uncertainty and reduces the reliability of the estimated PFS probability.

FDA performed an interval censoring analysis of PFS to assess the effect of a median 5-week PFS benefit relative to a 6-week imaging interval. Because tumor assessments occurred every 6 weeks, the exact date of disease progression is unknown and can occur anytime during the period between imaging assessments, as represented by the red shading in the patient follow-up timeline shown on the left of the slide. Since the median PFS difference of 5 weeks observed in CodeBreaK 200 was less than the imaging interval, the results are considered unreliable, as it cannot be ruled out that the difference is not due to inherent

measurement error.

Both FDA and the applicant performed an analysis of PFS using interval censoring to account for measurement error and timing of tumor progression assessments, as shown on the right. This analysis assumes that progressive disease events may have occurred at anytime during the imaging interval and not just at the end. The estimated median PFS results were 4.47 months for sotorasib and 4.3 months for docetaxel, with an estimated hazard ratio of 0.71. While the hazard ratio from this analysis is relatively consistent with the primary analysis result, the estimated difference in medians is approximately 5 days, which further adds to the uncertainty and the magnitude of PFS difference between the treatment arms.

In the setting of a primary PFS endpoint, FDA also evaluates overall survival, which is a more objective endpoint that provides important efficacy and safety information. In CodeBreaK 200, long-term follow-up for overall survival continues

to show no difference between arms. Relative to the sotorasib arm, there was a longer median OS and fewer deaths in the docetaxel arm. We believe this may be, in part, due to missingness of patient-level data, further highlighting the challenges and interpretation of the overall survival. Our FDA analyses show that the institution of crossover was unlikely to have any meaningful impact on the OS results.

In CodeBreaK 200, there were more deaths reported on the sotorasib arm relative to docetaxel. Our safety review did not identify any signals that explain the high rates of death in the sotorasib arm. Again, this may be due to high rates of dropout and missing data.

CodeBreaK 200 included secondary PRO endpoints for efficacy and tolerability. Although the statistical analysis plan included PRO endpoints in the hierarchical testing scheme, PRO endpoints were not formally tested because the test for overall survival indicated no difference. There were high rates of PRO instrument completion

by patients who remained on treatment, but this does not account for the asymmetric early dropout, and those patients were not offered the opportunity to respond to PROs.

Of the patients who received treatment, descriptive PRO information regarding side-effect bother demonstrated worst side-effect bother in the docetaxel arm. This supports the known toxicity profiles for both drugs. Interpretation of PROs is limited by a number of issues, including that there was no formal PRO comparison, the open-label design of the study, and the previously mentioned asymmetric early dropout. This result should be interpreted with caution, given that systemic bias can interfere with the interpretation of all endpoints, especially those with subjectivity and measurement such as PROs.

We will now discuss the findings of the FDA review of study conduct and potential systemic bias. FDA's review included an assessment of the confirmation of progression procedure, which revealed a potential study conduct issue. As

background, the applicant implemented a confirmation of progression, or COP procedure, at the time crossover was introduced to the trial. Rather than relying on the established blinded independent central review to confirm progression, which could take up to 10 business days, the COP procedure was implemented. This allowed separate radiologists from the BICR radiologists to provide a second opinion to investigators calling disease progression within 3 business days.

COP was required not only for crossover patients on the docetaxel arm to sotorasib, but also for patients who received treatment beyond progression on either arm; however, investigators would make the final treatment and patient management decisions. The potential impact of implementing this confirmation of progression procedure is usually minimal if it is used as intended.

Per the statistical analysis plan, an interim analysis for PFS was to be conducted at 70 percent information fraction. At the time of

this interim analysis, the PFS was statistically not significant and the independent data monitoring committee recommended that the study continue. As part of a separate process during periodic routine reviews to project the primary analysis timing, per the applicant's description, the applicant observed a higher than expected discrepancy between COP and BICR based events of progression. The applicant then raised concerns of this discrepancy with the imaging vendor.

The applicant used the COP procedure beyond the scope of its intended use when the applicant notified the imaging vendor of this discrepancy. The applicant's indirect input on the response assessments triggered a review process by the imaging vendor that led to a BICR re-read. This ultimately resulted in the identification of 12 additional PFS events, 11 from the docetaxel arm versus one from the sotorasib arm, leading to an updated PFS interim analysis that was statistically significant.

FDA considers these interactions a potential

violation of the imaging charter. FDA has attempted to elucidate further details from the applicant regarding these events. While the applicant has responded to all of our requests for information, FDA still lacks clarity regarding the interactions between the applicant and imaging vendor.

This potential misuse of the COP procedure resulted in an informal audit of the original BICR reads. The FDA views this as a potential study conduct issue. This also highlights the inter- and intra-reader variability of PFS assessments, which adds to the subjectivity of PFS as an endpoint. Ultimately, when the applicant presented FDA with these revised interim analysis results, FDA advised against the submission of a marketing application, based on the uncertainty surrounding the small PFS benefit over docetaxel, with only 12 new PFS events, changing the statistical significance and the uncertainty surrounding the re-reads. Given concerns of data quality, FDA expressed the importance of achieving consistency in BICR reads

from a single entity. Accordingly, the applicant elected to perform a global BICR re-read for the final PFS analysis.

You will now hear from Dr. Chuck Song, who will discuss three signals of potential systemic bias in CodeBreaK 200.

FDA Presentation - Chuck Song

DR. SONG: Good morning. My name is Dr. Chuck Song. I am the primary statistical reviewer for this application. As discussed earlier by Dr. Singh and Dr. Puthiamadathil, systemic bias is common in open-label trials such as CodeBreaK 200 because treatment assignment is known to patients and the investigators. Although bias is difficult to prove, data may signal its presence. It is noteworthy that not all signals of potential bias may result in bias in the efficacy estimation, but all introduce high uncertainty in the result and the study conduct.

For CodeBreaK 200, FDA identified three signals of potential bias. The first signal is the asymmetric early dropouts between treatment arms.

As presented earlier, there was an imbalance between trial arms in patients who were randomized but never treated. Twenty-three patients were never treated on studied therapy in the docetaxel arm compared to only 2 patients in the sotorasib arm. Most of these patients withdrew consent and were censored at day 1 for not having post-baseline assessments.

This imbalance suggests an open-label bias and the preference for treatment with sotorasib. This also has major implications for the statistical analysis, as early dropout predominantly on the control arm would lead to a loss of randomization.

So what is loss of randomization? We know that in a randomized clinical trial, the known and unknown prognostic factors are expected to be balanced by the randomization process. This is why randomized trials are considered the gold standard in evaluating drug efficacy, as the comparison between arms results in a treatment effect estimate that is fully attributable to the treatment of

interest. However, such balance will be lost if the patients who drop out are predominantly on one arm or are different from the other patients remaining in the trial. As a result, the trial arms would no longer be directly comparable and would introduce bias in estimating the treatment effect.

Although bias could be in either direction, depending on potential outcomes, given the incremental PFS benefit observed in this trial, FDA is particularly concerned of any potential bias that favors sotorasib. In other words, censoring of patients in the control arm will overestimate the PFS to the effect if these patients would have had better outcomes.

Our statistical review also found that the investigator imaging assessments appeared to favor the sotorasib arm. This signal was identified when examining discordances between investigator and the BICR assessment of disease progression. In this schema, patient follow-up is shown as the gray arrow and each vertical bar indicates an imaging

assessment. The red bars indicate investigator's call of progressive disease, while purple bars indicate BICR call. The assessment is concordant if the investigator and BICR determine progressive disease at the same assessment time.

Conversely, the FDA defines two types of discordant assessments. FDA defines early discordance as an investigator determination of progressive disease prior to the BICR and the late discordance as an investigator determination of progressive disease later than the BICR assessment. Overall, some discordance between investigator and the BICR assessment is expected in every trial and it does not necessarily indicate bias; however, when there is a differential distribution of discordance types across arms, this may signal the presence of systemic bias.

As shown in this table, among all of observed discordances in CodeBreaK 200, there is a higher proportion of early discordance in the docetaxel arm than in the sotorasib arm, and accordingly, the proportion of late discordance is

higher in the sotorasib arm than in the docetaxel arm. The difference is about 11 percent. This differential distribution of early and late discordances is suggestive of an investigator assessment of bias favoring sotorasib; in other words, these data suggested that either investigators were more likely to take patients off docetaxel earlier than they were to take patients off sotorasib, or they were more likely to keep patients on sotorasib longer than to keep patients on docetaxel, or some combination of both.

The third signal of potential bias in CodeBreaK 200 was the observation that patients in the docetaxel arm were crossed over to the sotorasib arm by investigators prior to BICR confirmation of progression. This aspect of the study design makes the primary endpoint of PFS by BICR vulnerable to the issue of informative censoring.

This schema depicts the relationship between investigator-assessed progressive disease and the BICR-assessed progressive disease in CodeBreaK 200.

The red bar indicates investigator call of progressive disease. The orange bar indicates when the patient would crossover to receive sotorasib, eligibility criteria were met for crossover, including the confirmation of progression by COP radiologists. Because there was no BICR call of progressive disease at the time of crossover, the BICR PFS of these patients would be censored at the last BICR assessment date.

In this example, this is shown as the dashed line at the time of the last investigator assessment. Censoring means that we know that the PFS per BICR assessment is at least as long as the solid part of the blue arrow shown in the figure, but its exact length is unknown because we do not know how long these patients' PFS would be after censoring, as shown by the hashed part of the blue arrow.

The follow-up for overall survival, on the other hand, is generally not affected by early crossover, as shown by the green arrow. Ultimately, although the BICR assessment is

performed by an entity, which is supposed to be blinded and independent, their assessments are not totally immune from study conduct issues, such as early crossover, based on investigator assessment, which may be subject to potential open-label bias. To be more specific, this raises a statistical concern of informative censoring.

The upper part of the schema on this slide is the same as the previous slide, showing patients censored for early crossover. The lower part of this slide depicts when patients crossover after both investigator and the BICR determined the progression. In CodeBreaK 200, we identified 19 patients who crossed over from docetaxel to sotorasib before disease progression was confirmed by BICR, resulting in censoring of their primary PFS endpoint. If these patients were healthier patients with better prognosis, their crossover would cause informative censoring, which in turn might have biased the results favoring sotorasib.

We compared overall survival after investigator call of progression for the

19 patients to the 27 patients who crossed over after BICR determined progression. The median OS was better with a lower event rate for the early crossover patients. An exploratory comparison of these two groups resulted in a higher ratio of 0.42 in favor of the early crossover group, indicating that patients censored due to early crossover may have had a better prognosis.

In summary, we have identified multiple signals of potential systemic bias in CodeBreaK 200. These signals generally decrease confidence in the observed results of the trial. Some of these signals could also manifest as statistical bias that impacts the estimation of the PFS treatment effect.

We now turn our attention to how the results of CodeBreaK 200 may differ from the observed results if patients who dropped out early or crossed over prior to BICR-assessed progressive disease were healthier than other patients in the docetaxel arm. In the following slides, we present a field of the sensitivity analyses performed by

FDA to characterize the treatment effect in the presence of the identified biases.

This is a tipping-point analysis for PFS about how the hazard ratio and the corresponding 95 percent confidence interval, represented by each dot and the bar, respectively, change with varying assumptions about the risk reduction of 20 early dropout patients censored for having no post-baseline assessment and the 19 early crossover patients.

For example, the left-hand vertical bar shows the PFS result. If we assume the patients with early dropout and early crossover are not different from other patients, still you follow-up, which is the primary analysis result. As we move right on the X-axis, we are gradually assuming a greater reduction in the risk of PFS events for the 39 early dropout and early crossover patients. From these results, if we assume the risk of a PFS event is 50 percent lower in these 39 patients, shown with the red arrow, the 95 percent confidence interval will include 1; in other words, the

statistical significance of the results would be lost. Based on the FDA analysis of the available data for early dropout and the early crossover patients, this appears to be a moderate and plausible violation of the non-informative censoring assumption.

We also examined whether the addition of crossover impacted the overall survival endpoint using sensitivity analysis. Different from the primary OS analysis, this sensitivity analysis attempts to estimate the treatment effect on overall survival under a hypothetical scenario in which no patient has crossed over. Ultimately, regardless of the assumptions made by this analysis, they all point to the same conclusion that there is no difference in overall survival across treatment arms in CodeBreaK 200. This analysis suggests that crossover is unlikely to be the reason for the observed lack of survival difference between sotorasib and the docetaxel arms.

In summary, the efficacy results of

CodeBreaK 200 are difficult to interpret because of the several signals of potential systemic bias. The potential systemic bias in CodeBreaK 200 may be difficult to overcome to reliably determine superiority of sotorasib over docetaxel, given the incremental PFS benefit and the no difference in OS. Finally, when addressing the statistical implications of the observed systemic bias, FDA's analysis suggests that the PFS benefit in Code Break 200 may not remain statistically significant if there is moderate violation of the statistical assumptions.

I now ask the cross-disciplinary team lead for this application, Dr. Paz Vellanki, to conclude our FDA remarks.

FDA Presentation - Paz Vellanki

DR. VELLANKI: Thank you, Dr. Song.

In CodeBreaK 200, sotorasib demonstrated an incremental PFS benefit and no difference in OS compared to docetaxel. The OS results were unlikely impacted by the 34 percent of patients on the docetaxel arm who crossed over to receive

sotorasib or received a KRAS G12C inhibitor as a subsequent therapy in a second-line refractory disease setting, demonstrating a survival benefit as a reasonable expectation for novel therapies.

Additionally, there were multiple signals of potential systemic bias in study conduct issues. While potential bias is present in many randomized trials in oncology, the efficacy results in CodeBreaK 200 were underwhelming and may not be sufficient to overcome uncertainty in the trial results. Our question to the advisory committee is whether we can reliably interpret and quantify the PFS improvement per BICR for sotorasib in the setting of potential systemic bias?

While PFS has been commonly used to support approvals in oncology, the PFS endpoint is inherently subject to bias. There was both intra- and inter-reader variability of PFS assessments in CodeBreaK 200. When the same BICR radiologist re-read imaging scans, new PFS events were identified, changing the PFS interim analysis results from not significant to statistically

significant. For all trials with primary PFS endpoints, FDA conducts sensitivity analyses to explore the strength of the primary analysis. We have shown that the magnitude of PFS benefit for sotorasib in CodeBreaK 200 may not withstand such sensitivity analyses.

While we often see asymmetric dropout in clinical trials, the magnitude of benefit may allow for robust statistical analysis and provide confidence in the effect of the drug in question; however, in CodeBreaK 200, the incremental PFS effect and lack of OS benefit made this more challenging.

The applicant asserts the results of CodeBreaK 200 are robust, as the PFS hazard ratio withstands multiple sensitivity analyses. FDA agrees the estimated PFS hazard ratio is generally consistent across those multiple analyses; however, the FDA tipping-point analysis showed the statistical significance of the hazard ratio may not hold under different assumptions regarding the level of informative censoring caused by early

dropouts and early crossover.

Additionally, a complete and balanced assessment of PFS also includes evaluation of the median benefit, event rates, and shape of the Kaplan-Meier curves. Both the applicant and FDA agree that based on an interval censoring method, the median PFS benefit could be as low as 5 days. We note the higher rate of PFS events on the sotorasib arm, though we acknowledge that this was in the setting of incomplete information with early dropout on the docetaxel arm.

The Kaplan-Meier curves showed a modest separation; however, given high levels of censoring, the latter half of the curve may not be reliable. This comprehensive assessment highlights uncertainty regarding the robustness of the PFS results and our ability to quantify the treatment effect of sotorasib.

There are multiple signals of potential systemic bias in CodeBreaK 200. There was a high number of patients on the docetaxel arm compared to patients on the sotorasib arm, who withdrew from

the trial once they knew of their treatment assignment. Investigator imaging assessments favored sotorasib and there was early crossover from the docetaxel arm to sotorasib before BICR confirmed disease progression. All of these individual patterns and behavior, when taken together, impact our ability to reliably estimate the primary PFS for BICR endpoint and the overall trial results.

The interpretation of PFS was impacted by a loss of information and investigative patient management. Differences in patient prognoses may have also allowed for overestimation of the PFS treatment effect. Importantly, there could have been many other impacts of the potential bias that are unknown and unmeasurable, including on patient selection, adverse event reporting, and patient-reported outcomes.

The applicant acknowledges the inherent risk of bias in CodeBreaK 200 as an open-label trial and implemented strategies to minimize bias; however, FDA is concerned that the mitigation strategies

were not sufficient to overcome the consistent trends in study conduct favoring sotorasib, which may have been influenced by bias, and because there's not a large improvement in PFS, interpretation of the CodeBreaK 200 study results remains challenging. Our analyses indicate a possibility that there may not be a statistically significant PFS benefit with sotorasib over docetaxel, and if there is, it is not reliably quantifiable.

Study design features are distinct from issues regarding study conduct. FDA takes an active role in providing feedback on drug development, including on-study design for clinical trials intended to support marketing applications. We did this for CodeBreaK 200. While features of CodeBreaK 200, such as the open-label design, may have increased susceptibility to issues with study conduct and potential bias, it is the responsibility of the applicant to both design and conduct trials, which can withstand and mitigate bias. In the case of CodeBreaK 200, a perceived

loss of equipoise, even prior to initiation of the trial, may have led patients and investigators to favor sotorasib overdose docetaxel, and led to a change in behaviors in the trial.

Public awareness for sotorasib, an oral drug against the previously undruggable target, which later demonstrated a moderate response rate in CodeBreaK 100, may have led to a perceived loss of equipoise in Code Break 200. It is possible that patients may have dropped out or withdrew consent to seek alternative trials evaluating KRAS G12C inhibitors. Patients may also have opted for standard-of-care therapy with their local oncologist to avoid the burden associated with being in a trial. CodeBreaK 200 highlights how patterns of behavior across multiple aspects of a trial may lead to concerns for potential systemic bias in favor of the investigational drug.

Moving forward, we hope to spend more time discussing how to mitigate bias in open-label trials. Potential strategies may include patient education to reduce withdrawal of consent;

investigator education to reduce bias related to imaging assessments; allowing for crossover to reduce dropout from the control arm; real-time BICR to reduce censoring related to discordant investigator and BICR assessments of disease progression; selection of an OS primary endpoint, which may be less impacted by potential systemic bias compared to PFS; and consent for OS follow-up, even if patients drop out of the trial, to maximize collection of data for a more reliable assessment of overall survival.

The FDA's regulatory considerations around CodeBreaK 200 take into account that the trial was conducted as part of the postmarketing requirement to verify the clinical benefit of sotorasib after the May 2021 accelerated approval, based on single-arm response rate data. When assessing whether the results of CodeBreaK 200 may be used to convert the accelerated approval of sotorasib to a traditional approval, we consider several factors, including but not limited to the following.

Can the PFS per BICR results be reliably

interpreted and can the magnitude of effect mitigate the uncertainty around interpretation of the primary endpoint? If so, then CodeBreaK 200 could potentially serve as confirmation of clinical benefit and fulfillment of the postmarketing requirements; however, if not, we would have an accelerated approval which is yet to be converted to a traditional or regular approval, and we would consider potential next steps within our regulatory framework.

After a confirmatory trial fails to verify clinical benefit, the regulatory decision to withdraw an accelerated approval is not automatic. The decision is affected by the overall results of the confirmatory trial. For example, a drug that demonstrates survival detriment may likely result in withdrawal of the accelerated approval. Another important consideration is the benefit-risk assessment in the context of the current treatment landscape rather than the benefit risk assessment at the time of the accelerated approval. A potential safety advantage of the drug over current

available therapy is also considered when deciding whether an accelerated approval should be withdrawn or whether there may be an alternative path to verify clinical benefit.

While sotorasib was the first KRAS G12C inhibitor to receive FDA approval, there are numerous competitor drugs currently being developed for non-small cell lung cancer. Adagrasib is the other KRAS G12C inhibitor farthest along in drug development, and it is the only other drug in class that has FDA approval to date. Adagrasib was granted accelerated approval in December of 2022 and the confirmatory randomized trial, KRYSTAL-12, is ongoing. KRYSTAL-12 evaluates the same patient population as CodeBreaK 200, has the same docetaxel control arm, allows for crossover, and also has a PFS per BICR primary endpoint. Per clinicaltrials.gov, the estimated primary completion date of KRYSTAL-12 is in May of 2025.

We note that the applicant has a planned randomized trial in the first-line setting. CodeBreaK 202 randomizes patients with KRAS G12C

mutations who are PD-L1 negative to sotorasib with chemotherapy versus pembrolizumab with chemotherapy. The primary endpoint is PFS per BICR. The results of this trial may be another potential way to verify the clinical benefit of sotorasib in lung cancer.

Given multiple regulatory pathways and the evolving therapeutic landscape, FDA is not seeking the advice of the advisory committee as to whether CodeBreaK 200 should be used to convert the accelerated approval to traditional approval for sotorasib, rather we are asking the committee to discuss the findings of the FDA review team, the multiple signals of potential bias, and if the observed PFS per BICR treatment effect can be reliably interpreted. We will use the committee discussion and conclusions to decide our next regulatory steps.

We would like the advisory committee to vote on the following question. Can we reliably interpret the PFS per BICR effect of sotorasib versus docetaxel in CodeBreaK 200? As a final

note, FDA recognizes the time and effort necessary to conduct cancer clinical trials. We would like to particularly thank the patients and their families, as well as the investigators and research staff who participated in the research studies discussed today. Thank you.

Clarifying Questions

DR. MADAN: Okay. I would like to thank this morning's presenters for staying on time, so we have our allotted one hour for discussion.

We will now take clarifying questions for both Amgen, Incorporated and the FDA. Please use the raise-hand icon to indicate that you have a question and remember to lower your hand by clicking the raise-hand icon again after you've asked your question. When acknowledged, please remember to state your name for the record before you speak and direct your question to a specific presenter, if you can. If you wish to have a specific slide displayed, please let us know the number of that slide if possible. Finally, it would be helpful to acknowledge the end of your

question with a thank you or end your follow-up with, "That is all for my question," so we can move on to the next question.

We will go through the raise-hand icon, which I think tells me who's first. Dr. Spratt, I believe, has the first question.

Dr. Spratt?

DR. SPRATT: Thank you. Dan Spratt, Case Western. Thank you both for all the work put into this. It's three interrelated questions, and I'll make them concise. This is for Amgen.

It's in your briefing document, table 11 or figure 15. If we believe your PFS-1, your primary endpoint, is superior for your drug and your data on PFS-2, the effect size estimate -- also favored although not statistically significant -- was also superior for your drug, the question is, why would overall survival be similar or potentially worse? So that's question one, and that leads to can we reliably interpret your PFS results?

Question two is, you kindly did report --

DR. MADAN: Maybe, Dr. Spratt, we'll let

them answer question one, and then that way, they'll be able to remember question two.

If the sponsor could address question one from Dr. Spratt? Thank you.

DR. FRIBERG: Yes. Thank you, Dr. Spratt, for the question. The purpose of performing the post hoc PFS-2 analysis was to put the overall survival results into context. One of the questions that logically comes up when you see overlapping Kaplan-Meier curves was, was there something that happened after progression that led to a detriment in the next line of therapy? That does not appear to be the case. It does not explain why the OS benefits were similar in the two arms.

DR. SPRATT: Okay. Thank you.

That would lead me to believe that if PFS-1, and potentially PFS-2, by the way they were measured, were superior with the drug, we still don't then have a clear answer why OS would be no difference.

The second question is also for the sponsor.

The restricted mean survival times that were shown in table 18, it's about a 1-to-1-and-a-half months restricted mean survival time benefit for PFS. It was unclear to me. The overall survival Kaplan-Meier curves as they crossed, I don't know if you tested -- did this violate proportional hazards? And I didn't see the restricted mean survival time for overall survival. Has that been performed?

DR. FRIBERG: Yes, that has been performed. I'd like to take the opportunity to ask Dr. Koch to take the podium and respond to your question regarding this.

DR. KOCH: I'm Gary Koch, professor of biostatistics, University of North Carolina, Chapel Hill. My institution is compensated for my time, and neither I nor my institution have any financial interest in the outcome of the meeting.

If we look at slide 1, this reports the RMST results for PFS and it shows that at 12 months, the difference is 1.33 months and at 14 months -- this is from follow-up from baseline to 14

months -- it's 1.61, and this represents a difference in means of survival over those intervals. One can additionally have some interpretation by dividing the difference by the length of follow-up. So if we divide 1.61 by 14, we're viewing the area between the Kaplan-Meier curves, which is what the difference in RMST manages, as like a rectangle, and the 11.5 percent in the right-hand column means that the average difference in PFS rates over the 14 months is about 11-and-a-half percent.

In slide 2, we essentially have corresponding results for overall survival, and on overall survival over 24 months, the difference in the means is 0.17, and then again, if we look at the difference over the 24 months by dividing the 0.17 by 24, we get a confidence interval from minus 7.1 percent to 8.5 percent, with the 7.1 percent being the difference in favor of docetaxel, and that's the amount of difference that might possibly be ruled out by the difference in RMSTs.

With respect to departures from proportional hazards, when the curves are on top of one another, then typically there would not be any difference between proportional hazards over the follow-up time.

DR. SPRATT: Thank you so much.

If people are able, the last one is on your slide CC-32. I didn't hear anyone comment. It does appear you had 7 hyperprogressors on your far left. I didn't know if that was something relevant to comment on.

DR. FRIBERG: Yes. I'd like to ask Dr. Mehta to comment on the progressors in each arm.

DR. MEHTA: Thank you. Slide 2, please? We did look at these patients on the red bars in the waterfall plot, and this slide here shows the two arms and the numbers of patients that were in the red bars. You had 10 patients on the sotorasib arm and 12 patients on the docetaxel arm, whose best response was progressive disease. And we looked at these patients' greater details, specifically the

3 patients for whom you see the spikes, so to speak, in the disease, and all three of these patients had low tumor burden to begin with, and these spikes represent a relative increase in the tumor size and not a hyperprogression as such.

We also looked at the molecular characteristics of these patients in the red bars, and there appeared to be no significant enrichment of any co-alterations for the small set of patients whose best response was an increase in tumor size of greater than 20 percent. Thank you.

DR. SPRATT: Thank you. That's it for me. I really appreciate it.

DR. MADAN: Okay. Great. Thank you.

Our next question is from Dr. Vasan.

DR. VASAN: Hi. Thank you to both the FDA and the applicant for this really careful analyses. I had two questions. One is these tipping-point analyses, because it seems to me that that is a source of discordance between the applicant and the FDA, so for the FDA, this is slide number 36, and for the applicant, this is slide CC-87.

It seems to me that the FDA's analysis, the X-axis, is this percent risk reduction, so binning patients together, whereas the applicant's analysis is sort of this patient-by-patient analysis. Obviously, I think the interpretation of these two analyses is quite different. So I was wondering if both the FDA and the applicant could comment on the merits of the way that they analyzed these data and why that advocates for their position.

DR. SINGH: Thank you, Dr. Vasan, for the question. We would like to invite the applicant to respond first, and then I will invite Dr. Chuck Song to comment.

I just want to add that in terms of the term "statistical pessimism," which I think is something you may be alluding to, which the applicant used this term a few times, and you'll hear this from Dr. Song, I want to say that the role of the FDA actually is to make conservative and moderate estimates of assumptions, statistical assumptions, because certainly we would not expect individual drug sponsors to perform those types of analyses.

So at this moment, I'll defer to the applicant, and then invite Dr. Chuck Song to respond.

DR. FRIBERG: Thank you for the opportunity. This is a critical point. The analysis used similar methods, as you point out. What the agency refers to as some moderate statistical methods, I think we've already described are clinically actually quite extreme when you look at what it means to be a 50 percent improvement in PFS.

That being said, your question is a bit more of a philosophic one and, again, I think Dr. Koch is well positioned to be able to answer this question.

Dr. Koch?

DR. KOCH: Gary Koch, University of North Carolina, Chapel Hill, statistics. There were different types of sensitivity analyses that were produced by the sponsor. One of them that was described dealt with the early dropout or discontinuation from the study by the 23 docetaxel patients prior to being treated right after

randomization.

The analysis the sponsor did in that particular case was to impute outcome for them, reasonably optimistically, from the patients with at least 6 weeks of follow-up, and the sampling to do that is shown in slide 1, where basically the 23 patients were repeatedly randomly sampled from the patients with follow-up at least 6 weeks, and the results of that analysis were then shown in slide 1 again, CC-74, so basically that was very supportive. More pessimistically, the sponsor also did such an analysis by randomly selecting from the patients with at least 12 weeks of follow-up, and that analysis was similarly supportive. These analyses were definitely favorable to the docetaxel group by assuming reasonably optimistic results for those particular patients.

Now, the difficulty with the analysis referred to in the presentation from FDA, as described in slide 2, is that that analysis more optimistically did the selection from the best 50 percent of patients, and in particular, it's

noted that those patients would have essentially a fairly favorable median, although when it was from the docetaxel group, the analysis was still reconfirmed for the original primary.

Now, the sponsor did a second type of analysis, which was concerned with the patients who had early censoring due to basically crossover to other treatments, and that was initially reviewed by Dr. Friberg in slide 1, where the most pessimistic possible paradigm was assumed for the sotorasib patients by basically assuming that all 24 would have been a BICR event, essentially at the time of starting their new therapy.

Then for the 31 docetaxel patients, what was then done was to assume that none of them had a progression at the time of the start of new therapy. And there, as you see in red, the confidence interval just barely crosses 1, but if one is willing to say at least one of them would have had an event, more or less, at the time of the start of early treatment, then the results would have then become favorable.

The sponsor then additionally in slide 2 put the patients with getting a new anti-cancer treatment and were censored for that reason, and that was the 24 and 31, with the 2 and the 23, and there, again, one only needed to see three progressions on the docetaxel arm on these particular patients in order to restore the result in favor of sotorasib.

These are the kinds of analyses the sponsor did. The one that you see in this slide is very pessimistic to sotorasib by assuming that all 26 of these patients would have had events, while only a minimum number of them with docetaxel need to have an event at the indicated time in order to restore the original result of a positive result for PFS for sotorasib.

DR. SINGH: Thank you. I'd like to invite Dr. Chuck Song from FDA to respond briefly to this.

DR. SONG: Thank you for the question, and also, thank you to the sponsor for discussing the difference between sensitivity analysis. I first want to reiterate what Dr. Singh just said, because

our role as FDA reviewers, we must consider a more conservative view of the data than the sponsor, and second, I want to clarify our analysis.

Could you please show the FDA slide of the tipping-point analysis?

FEMALE VOICE: What number slide, please?

DR. SONG: It's slide number 33, I think. Can you go to that slide, number 23?

FEMALE VOICE: Is this the main slide deck or --

DR. SONG: Main slide.

FEMALE VOICE: Thank you.

(Crosstalk.)

DR. SINGH: It's slide 36 in the main slide deck. Apologies for the confusion.

DR. SONG: Okay.

DR. SINGH: You had it up a moment ago.

DR. SONG: Okay. So this is our analysis, and this is not the analysis that the sponsor criticized for being too conservative. That analysis is a supplement analysis, which we impute based on the top 50 percent of patients. But in

this analysis, we didn't impute based on the top patients, but we're saying after these patients are being censored, the risk of an event will reduce by a different percentage, ranging from 0 to 90 percent. We found at 50 percent reduction of the risk, the confidence interval will cross 1. So this is a different analysis that I want to clarify.

Could you also pull up our backup slide number 29, backup slide number 29? I want to address the sponsor's sensitivity analysis because in the sponsor's sensitivity analysis, they talk about these 24 patients censored for new anti-cancer therapy in sotorasib and 31 patients censored for new anti-cancer therapy in the docetaxel arm. They treated them all as events, or they treat all of the sotorasib patients as events and the docetaxel patients as non-events.

We actually looked into the overall survival of these two groups of patients, and you can see that the median overall survival for the 24 patients was 11.2 months and for the

31 docetaxel patients censored for new anti-cancer therapy, they survived 7.4 months longer. Also, if you look at the survival post-censoring for new anti-cancer therapy, the difference is still there. The docetaxel patients censored for new anti-cancer therapy survived 6.6 months longer, and the hazard ratio analysis, also pointing, the sotorasib patients who got censored for new anti-cancer therapy are more unhealthy.

Can you go to the next slide? As we said, among these docetaxel patients censored for new anti-cancer therapy, 19 out of them were actually crossover patients, and we have already shown these 19 patients had very good overall survival. They survived 24 months in terms of median, and post-censoring, they survived 17.7 months.

The next slide please? So we actually did a similar analysis as the sponsor did, and the second row of this analysis, we treat the new anti-cancer therapy in sotorasib only as an event, and we got the same result as the sponsor's tipping-point analysis. But if we treat the new anti-cancer

therapy as an event, except for the early crossover, the 19 patients, you can see the results still getting more towards 1 and the upper bound of the confidence interval becomes 0.94. and highlights the uncertainty of the data.

DR. SINGH: Okay. Thank you, Dr. Song. I think that adequately addresses the question. Thank you.

DR. MADAN: Okay. Thanks.

I think we've got several questions lined up, so I'll ask each questioner to give their most important question, and then move on to the next one so everyone has a chance. And I'll ask responders to be direct and on point to the question so we can get all these discussion points in.

Our next question will be from Dr. Nieva.

DR. NIEVA: Thank you. My question is for the applicant. I'm Jorge Nieva from USC. My question is regarding the blinded independent central review. I'd like to know what was the nature of the errors in the first BICR analysis?

I want to know if the first vendor acknowledged that there was some kind of incompetence in their analysis, because I'm concerned that basically there were two chances to hit on PFS by doing the analysis twice, and that may have bias. I'd also like to know if the COP analysis that was done, that differed from the first BICR analysis, was informed by the opinion of the treating physicians and if there was communication between the two. Thank you. That concludes my questions.

DR. FRIBERG: Thank you for the question. Before I bring up any slides, I just want to be unambiguous about three points. One, we did not violate the imaging charter. Number two, the blind was maintained with regard to treatment assignment at all times on the study. And number three, again, even if you believe that there were challenges, the 100 percent re-read should reset and nullify those concerns.

That being said, if we could bring up slide 1, and I also want to clarify that the FDA

scheme -- I think it's figure 2 -- which they also showed in their analysis, is actually not quite correct. The scan data is shared with investigators, as well as the COP assessment and the BICR. The investigator was never a gatekeeper, per se, in order for the BICR to receive information. In that regard, again, they were blinded to treatment assignment at all times, and there was no communication between the investigators, the COP assessment, and any of the BICR assessments.

With regard to -- I think you had a third question. Can you repeat your third question that was in embedded in there?

DR. NIEVA: Yes. I'd like to know the nature -- we do blinded independent central review because we presume it to be more competent or informed. So the fact that the first blinded independent central review seemed to have a large number of errors is concerning. So I'd like to know if that's something that's been acknowledged by the vendor or if the vendor stands by their

initial assessments, and I'd like to know if there was a systematic nature to the types of errors that were being made.

DR. FRIBERG: So to clarify, the BICR process and the independent imaging reads were entirely independent. As I mentioned, no information was sharing. Also to put it into some context, less than 10 percent of the total reads that were performed by the BICR went through the COP process.

That being said, this aggregate data that was identified as having some discordances through the mechanisms that we described was through routine and outlined in the imaging charter communications with the imaging vendor. That led to an independent quality review at the level of the imaging vendor, and ultimately that led to them independently, without regard to saying which of the individual scans were involved; or without, again, knowledge of the treatment assignments, that led to their independent evaluation, reader retraining, and ultimately the three scans that had

their values changed.

So in that regard, the auditing that the FDA brings up would only have been possible through this communication with Amgen that, again, was without regard to treatment assignment, and the global 100 percent re-reads should have nullified that. So again, no imaging charter violation and the re-reads should have accounted for all of this.

DR. MADAN: Okay. Thanks.

DR. SINGH: Dr. Madan, may I be permitted to just respond, since it was basically said that the FDA is being inaccurate? I think that we did say within our presentation that this was a very confusing process for us to elucidate. We called it a potential violation, and we did try to gain a deeper understanding. Nevertheless, we considered this to be, in totality, just an atypical interaction, triggering a series of re-reads, which again speaks to just the global concerns regarding the fidelity of this endpoint. I'll end there.

DR. MADAN: Thank you, Dr. Singh.

Dr. Shaw, you have the next question.

DR. SHAW: Thank you very much. Pamela Shaw, Kaiser Permanente, Washington Health Research Institute. I just had a couple of quick follow-up questions regarding the BICR re-read process. I just wanted to understand, were those completely new people from that vendor or new organization that were re-reading it -- so that would be the first time they saw the scans -- or was it some of the same people reading the same scans a second time?

DR. FRIBERG: Thank you, Dr. Shaw. They were three separate new individuals, new radiologists, who were independent from anyone who had ever seen a scan on the study.

DR. SHAW: Okay. Great. Thank you. I think that completes my questions about the BICR.

Then I just had another second question, which related to understanding some of these sensitivity analyses, and we've heard the term "pessimism" being used in some of those imputations, where we think about those people that stop treatment or the early crossovers, and we

heard about this imputation process where we take the 58 percent, the top 50 percent, in terms of best progression events in the imputations.

For me, what I understood -- and maybe this is a question for Dr. Song, and you can tell me if I'm interpreting this correctly -- is that we've learned about the people, particularly, I'm going to call it the doxa [ph] arm -- I don't pronounce it very well -- that the early switchers had better survival. I think it was a 42 percent hazard ratio. And also, there was a differential better survival being censored for the standard arm.

So the idea that this 50 percent imputation is optimistic, I'm confused because the way I think about it, if I'm going to impute this progression, I want to think about people with a similar prognosis. So I'm actually wondering, rather than just taking the progression times, did you think about doing an imputation, or did anyone do an imputation, where you think about people with similar prognosis, similar survival, and then look at the progression times, the progression-free

survival times, amongst those who had, obviously, better survival that we're getting censored on one arm versus the other?

I don't know if that question made sense, because I'm not sure if an optimistic implication was done because the survival wasn't considered, and it seemed like there was a survival difference or at least some evidence of that.

DR. MADAN: Maybe I can try to distill that, and you can correct me if I'm wrong, Dr. Shaw. But you're asking, basically, with the statistical extrapolations, were they done with patients of similar characteristics so you could have a more accurate imputation?

DR. SHAW: Yes, in terms of the prognosis, because I'm concerned that this term "optimism" is giving us all comfort, and I'm not sure they were optimistic at all because they didn't consider one of the most important characteristics of the patient, which was prognosis, and somehow conditionally imputing on prognosis, based on survival times.

DR. MADAN: Yes. And can you just clarify who you're asking the question to?

DR. SHAW: I'll ask Dr. Song, first, whether he considered an imputation like that, and then I'd be happy to hear from the Amgen group because they did a lot of thoughtful sensitivity analyses as well.

DR. MADAN: Thank you.

Dr. Song?

DR. SONG: Yes. Hi. Yes. We did imputation analysis for this group of patients because we all deserve that they have better survival for their overall survival, so we assume they have better outcome for the PFS also. But the survival, we didn't know the missing part of the PFS, so we cannot really -- because this is a missing data problem.

DR. SHAW: I see what you're saying, yes. Okay. Thank you for that. Thank you.

DR. MADAN: So I guess the sponsor, would you guys like to reply?

DR. FRIBERG: Yes. Thank you for the

opportunity. I'm going to ask Dr. Suresh to come up and comment. I think we have some additional data that could be helpful here, both with what the Kaplan-Meier curve looks like for a 50 percent lower event rate from the tipping-point analysis, as well as more broadly about the wide variety of sensitivity analyses that we performed.

DR. SURESH: Ram Suresh, oncology, biostatistics, Amgen. To answer Dr. Shaw's question, bring up BU-320, please. First, let me talk about what we did for the 23 docetaxel patients who dropped out. What we did was -- slide 3, please.

DR. MADAN: To clarify, we're not seeing slide 23. Okay. Thank you.

DR. SURESH: Okay. So what we did was Dr. Friberg showed the sampling where we made an attempt to sample from enriched patients who survived at least 6 weeks. In other words, we excluded all the early progressors, and their deaths, and the censoring, and the 120 patients that were enriched from which we sampled.

By doing so, this was an attempt to try to get an answer for the unobserved variants because we are sampling from the same docetaxel pool. Additionally, we wanted to include a degree of stress, and then went ahead and sampled for these 23 patients from all the docetaxel patients who had not progressed, died, or censored by 12 weeks. And when we did this, as is shown on the screen, the hazard ratio is 0.73, and 83.9 percent of the times, the results were statistically significant. I just wanted to submit this.

DR. SINGH: Great. Thank you.

DR. FRIBERG: If I could just bring up one additional slide, slide number 2, I think we've been talking about how extreme are some of these assumptions, and this is an image that shows, again, the progression-free survival estimates of what actual docetaxel patients from the study are. And you see that, again, the original, and it's a grayish brown here, and the light blue represents a 50 percent risk reduction from the original.

DR. SHAW: Okay. That's really helpful.

And just a quick follow-up, the FDA seemed to note the survival difference, not just for those that did the early crossover, but also those who may have discontinued due to AEs. When you did that particular imputation you're referring to, did it include that expanded group who discontinued due to AEs or just the early crossovers?

DR. FRIBERG: That particular simulation I'm going to have Dr. Suresh comment on.

DR. SURESH: In our simulation, the sample is from patients who had not progressed, died, or censored until 12 weeks, and there is evidence that they are continuing beyond 12 weeks.

DR. SHAW: Okay.

DR. SURESH: Can I give you another perspective also related --

DR. SINGH: Well, I believe that the FDA would like to respond to just a few of these assertions very quickly, and I would ask that Dr. Pallavi Mishra-Kalyani quickly responds.

DR. MISHRA-KALYANI: Sure. This is Pallavi Mishra-Kalyani, FDA, statistics. First, Dr. Shaw,

thank you very much for your question. I think it was an excellent one, and certainly imputing PFS based off of knowledge of prognosis on survival time is a good one, but as my colleague, Dr. Song, mentioned, it's very difficult with missing data problems to be able to identify the group, the correct group of patients for imputation, particularly given that there are so few patients, in general, with long survival times in this study.

Secondly, I think you mentioned the two groups of patients who dropped out, or had early crossover and also dropped out, and the dropout patients that we are describing mostly didn't even receive a single dose of therapy, so they weren't necessarily dropping out early due to adverse events; they were just dropping out very early into this study after randomization.

Lastly, I'll just mention that with the analyses described by Amgen, we don't disagree with their analyses, but these are very, very mild assumptions about these patients, and we've already noted from our additional analysis that these

patients do tend to have better prognosis. So it's better to take a more moderate approach in assuming that they would have better PFS had they stayed on study and then observed with the BICR PFS. Thank you.

DR. SINGH: Just a note, I would just like to say that we have a heavily clinical committee who is trying to understand complex statistical concepts, and I think the understanding is that both the FDA and the applicant performed various sensitivity analyses to interpret the robustness of the results, and different assumptions can be used with different results, again, highlighting our overall challenges in interpreting these trial results.

DR. SHAW: Yes. I would like to say that I feel like this has been a great discussion, and I agree there's been very reasonable sensitivity analyses done on the part of the sponsor and the FDA. And I think my main conclusion, or the reason behind my question, is I just caution the use of pessimistic because it is difficult to understand

what would be pessimistic versus optimistic.

DR. MADAN: Thank you, Dr. Shaw. We're going to move on to the next question --

DR. SHAW: Thank you very much.

DR. MADAN: -- and return later if we can get through the questions. We still have panel members who are waiting patiently.

Dr. Hoffman, you can ask your primary question and direct it to either the sponsor or the FDA.

DR. HOFFMAN: Yes. I have two related clinical questions, probably best for Dr. Mehta. In view of the fact that I believe 99 percent of the accrual to CodeBreaK 200 had occurred prior to the accelerated approval date for sotorasib in 2021, with that number of people that dropped off the day after randomization, if you will, if they were not happy with having been randomized to docetaxel, was there an option at that time? Was there like an expanded access trial or something where they knew they could get sotorasib some other way? Was that the main issue? And then I'll ask a

short follow-up after that.

DR. MADAN: Okay.

DR. MEHTA: Thank you for your question. No, there was no other non-trial access to sotorasib at that time point. We did have expanded access programs ongoing; however, the expanded access programs had a clear eligibility criteria that they could not have received sotorasib on trial, and only if they had some eligibility limitations for the trial would they be allowed to access the expanded access protocols.

I also want to note that while you are accurate that the vast majority of patients had already been enrolled at the time of crossover amendment, the actual number of progression events was only 25 percent. So only 25 percent of BICR PDs had occurred when protocol amendment 3 was implemented, and that means about half of the docetaxel patients, or approximately 50 percent of the docetaxel patients, were able to access the crossover because by the time they experienced PD, their site had implemented the amendment and,

hence, were able to access crossover.

DR. MADAN: Okay. Thank you. I think that answers the question initially.

Dr. Hoffman, your second question?

DR. HOFFMAN: Yes. I wondered whether there was some at least general information about what the response rate was to those people who did crossover to sotorasib from docetaxel, and perhaps vice versa, if that information is there.

DR. MEHTA: Certainly. We can show you the outcomes of the crossover patients. Slide 1, please. On this slide, you see the swimmers plot on the left, but first I draw your attention to the table on the right. These are the 46 patients who crossed over from docetaxel to sotorasib. They came from numerous sites. Their median time on sotorasib after crossover was 4.8 months. Of these 46 patients, 10 experienced response, so that's ORR of 21 percent, and the disease control rate was approximately 76 percent.

We do not have a median OS that was achieved at the time of the data cutoff on these crossover

patients. The left panel is the swimmers plot. If you have additional questions, I can walk us through the swimmers plot as well.

DR. HOFFMAN: No, that's fine. Thank you.

DR. MADAN: Alright. Thank you.

DR. MEHTA: Thank you.

DR. MADAN: We'll move to our next question from Dr. Gulley.

DR. GULLEY: Yes. Thank you so much. This question is for the applicant. For those patients who crossed over early in the docetaxel arm before the BICR PD based on COP reads, what was the average RECIST percent increase or decrease in the final BICR reads?

DR. FRIBERG: That is a question I'd like to ask Dr. Mehta to come address.

DR. MEHTA: I understand your question to be the change in the RECIST target lesion size at the time of BICR PD --

DR. GULLEY: At the time of --

DR. MEHTA: -- for the crossover patients?

DR. GULLEY: -- COP, at the time at

crossover and they didn't have the BICR PD, yes.

DR. MEHTA: Okay. We do not have that information collated on the slide, but we can try to get it to you before the end of the day.

DR. GULLEY: Okay. Thank you.

DR. MEHTA: Thank you.

DR. MADAN: Thank you very much.

Okay. Dr. Rosko?

DR. ROSKO: Hi there. Ashley Rosko, Ohio State. My question is for the FDA. It's in regard to the frequency of when a new and independent BICR review is requested, this really speaks to the potential misuse of the COP procedure in which there was an atypical BICR re-read resulting in the development of a new PFS.

To the FDA, is there a threshold of discordance between the investigator assessment and a BICR assessment, and would you recommend a new and independent BICR team?

DR. SINGH: Thank you, Dr. Rosko. I'll start, and I'll invite Dr. Amatya to join. We noted in our presentation -- and I think your

question is, is there a threshold, basically, for which we recommend a global re-read? I can tell you upfront there is not. We typically see discordance between investigators and blinded central reviews; that's why we have blinded central reviews. That rate of discordance is typically about 30 percent.

What was concerning here was the direction of the discordance, the proportionality, the same bias illustrated on both arms, so calling progression earlier for patients on the docetaxel arm, then the blinded readers, and similarly calling progression later on the sotorasib arm, then the blinded readers, both patterns of behavior suggesting this implicit bias towards sotorasib.

When we recommended the global re-read at the time of the interim analysis, which again was narrowly flipped from negative to positive based on 12 patients, 11 of which were on the docetaxel arm, this was not triggered by some sort of threshold; it was triggered by just a concern, again, around the integrity of this endpoint.

Dr. Amatya is our biostatistician, and he can comment briefly on some more background on this.

DR. AMATYA: Thank you, Dr. Singh.

I think Dr. Singh covered and answered your question. This really was triggered by observed discordances between COP read and BICR read, and not because of any particular threshold. Thank you.

DR. MADAN: Dr. Rosko, does that complete your question?

DR. ROSKO: Yes. Thank you.

DR. MADAN: Thank you. I'm just making sure.

Our next question will be from Dr. Conaway.

DR. CONAWAY: Yes. Mark Conaway, University of Virginia. Thank you. My question is for Dr. Friberg, and we're back to slide CC-72. My apologies for going back to a slide we talked about a lot. You imputed from this pool in the green box at the bottom. If you plotted the PFS experience of those patients in that pool on CC-78, what would

that curve look like?

DR. FRIBERG: I'd like to ask Dr. Suresh to come up and present that data.

DR. SURESH: Can you bring up BU-636, please, slide 1? This is the display of the curve for 6 weeks. And could you bring up the 12 weeks? It's BU-637. Okay. Slide 1, please. This is for the 12 weeks, and I hope I answered your question.

DR. CONAWAY: Yes. Thank you. We'd expect they'd be shifted to the right. I was just trying to get a sense of how much, so thank you.

DR. SURESH: Thank you.

DR. MADAN: Okay. Alright. I think I'll ask a question, and we have other questions. But if you want to get back in the queue, there's still a little time before lunch here.

This is Ravi Robbie Madan from the National Cancer Institute, and my question is to Amgen. We've had a lot of discussion about the different statistical permutations, but a large part of our conversation has to do with the perceived minimal benefit in terms of PFS. What were the statistical

presumptions about a median difference in the initial design of the trial? In other words, what was the expectation for benefit, and was that expressed in a median benefit timeline?

DR. FRIBERG: The study was designed, rather than based on medians, it was designed on a relative risk reduction. And whether we're talking about the initial protocol or after amendment 3, it was always held stable at a 35 percent relative risk reduction; said another way, looking for a PFS hazard ratio of of 0.65, and ultimately what we observed was a relative risk reduction of 34 percent.

DR. MADAN: Thank you for that.

DR. AMATYA: I'll just respond if it's ok.

DR. MADAN: Go ahead. FDA wants to respond.

DR. AMATYA: This is Anup Amatya from the FDA. Initially, it was designed to detect 3.2 months of median difference.

DR. FRIBERG: The initial assumptions were before there was any data available for real-world evidence related to docetaxel performance on G12C,

as well as our ultimate results from CodeBreaK 100. It's fair to say that those were optimistic assumptions.

DR. MADAN: Well, I'm sorry. Just to clarify, there was mention of an expected result of 3.12 months, so why should we not hold that in regard here? I'm sorry.

DR. FRIBERG: So the medians, of course, are derived from the relative risk reduction, and whatever you plug into them, you can look at that difference. The assumptions there that read through, I think they were quoted here, but those were hypothetical. The relative risk reduction, which looks at, again, the entirety of the curve, was held stable between the different amendments, and ultimately turned out to be what we had predicted, or at least roughly, at the minus 34 percent relative risk reduction. Medians are only one way to show the result.

DR. MADAN: Okay. Thank you.

I believe Dr. Pantelas is next.

MR. PANTELAS: I appreciate the promotion,

but I'm not a doctor; I'm a patient. The question that I have is about a loss of equipoise, and the problem that I have is that when this drug was talked about at ASCO and at IASLC, it hit the KRAS community pretty hard. I don't know that you can get a lack of bias within the patient community, especially this kind of community, even if you're looking at a non-superiority.

Docetaxel is not seen as a kind treatment in this community, and it has very visible side effects. So creating an oral alternative to doxy, and one that has more patient friendly side effects, it creates a desire in the community for for noninferiority. And if you mention noninferiority with an oral option versus an infusion option, I just don't know how you take that into context in creating this trial.

DR. MADAN: Okay. Mr. Pantelas, if you take your perspective -- I guess I'm trying to understand your question to either the sponsor or the FDA. Is it one about the noninferiority interpretation of this data? Is that what you're

asking?

MR. PANTELAS: Well, whether or not the noninferiority interpretation of the data has value and supports continuation.

DR. MADAN: Alright. I think we'll start with the FDA, and then Amgen will have a chance to respond about this data and whether it supports noninferiority.

DR. SINGH: Okay. I see Dr. Amatya and I think Dr. -- well, I'll invite Dr. Amatya to comment on the noninferiority and I'd like to make a comment subsequently regarding the comments surrounding equipoise.

Dr. Amatya?

DR. AMATYA: Yes. Commenting regarding noninferiority, first of all, lack of superiority does not does not necessarily mean noninferior. What this data suggests is that there is no evidence of superiority. It doesn't necessarily show; it's no difference statistically. So I would rather caution against interpreting this as a noninferior result.

DR. MADAN: Thank you.

Dr. Singh, you can address the equipoise issue, then the sponsor to respond as well.

DR. SINGH: Yes. Dr. Puthiamadathil, would you like to address equipoise and mitigation strategies?

DR. PUTHIAMADATHIL: Thank you, Dr. Singh, and thank you, Mr. Pantelas for that question and comment. We agree with you. There was a significant amount of press ahead of time, so we do believe that this actually impacted patient perspective, as well as investigator perspective. Dr. Johnson in her presentation actually said that physicians love to hate docetaxel and the patients dread it, so she's not surprised by the early dropout. So that really indicates the sort of milieu that this trial was going through. There was what we believe was perceived as loss of equipoise.

If you can go to slide 47 in our main presentation, we can discuss some of the mitigating factors that are available to us that can help

potentially address these issues of perceived loss of equipoise. Obviously, these include patient and investigator education. I think this certainly could have done a little bit better, I think, in terms of patient and investigator understanding about where sotorasib really stood compared to docetaxel. We could also obviously have had -- the study sponsor did increment crossover.

Real-time BICR assessments also help in terms of getting patients to determine when there is real progression, and also, obviously, the endpoint selection. We've discussed how PFS is inherently subject to bias versus overall survival, which is a more objective endpoint, and obviously, for the long term, we can suggest collection of OS follow-up even when patients withdraw early. Thank you.

DR. SINGH: Thank you.

DR. MADAN: Thank you very much.

So we are coming up on lunch. Dr. Kraus has one question --

DR. FRIBERG: Could I --

DR. MADAN: -- but I wanted to give Amgen a chance to balance out the discussion here if they have any comments on either the noninferiority interpretation of the data or the equipoise issue.

DR. FRIBERG: As the sponsor, can we reply to that as well? Could it be possible?

DR. MADAN: That's what I just said.

DR. FRIBERG: Oh, I'm sorry. I thought you were asking other members of the committee.

DR. MADAN: [Indiscernible] to address either the noninferiority question interpretation or the equipoise question.

DR. FRIBERG: Thank you so much. One thing that I think I want to make sure we're not losing track of is this noninferiority discussion about the overall survival. The PFS result was statistically superior and, again, we've looked at a variety of tests. To address this directly, though, I do think rather than going into additional methods or additional statistics, I think asking Dr. Johnson to comment a bit on what that means and, again, her perspective would be

appropriate.

DR. JOHNSON: Thanks Dr. Friberg.

Thanks for the opportunity. I do think it's critical -- I'll piggyback on what Mr. Pantelas commented upon, which is while there was buzz in the community, I think to blame early dropouts and to attribute it all to perceived loss of equipoise is short-sighted. It's not subtle what docetaxel does to fragile patients.

So while we can look at the 20 patients that decided not to enroll in the study -- Dr. Friberg showed us a nice analysis about how those patients were actually less fit or even sicker than than the larger group. But for any patient that received docetaxel, as a clinical oncologist, we know what happens, as those patients decline quickly, and that results in constitutional symptoms; that results in patients appearing as though they they are progressing and declining in performance status. So that point hadn't come up yet, and I think to just call this statistical noise would be a shame.

DR. MADAN: Thank you very much.

Our last question before lunch will be from Dr. Kraus.

DR. SINGH: Dr. Madan, Dr. Pazdur has joined and he has a comment.

DR. PAZDUR: I want to address that comment.

(Crosstalk.)

DR. SINGH: May we allow this comment?

DR. PAZDUR: Yes. I want to address that comment.

DR. MADAN: Yes. Go ahead, Dr. Pazdur.

(No audible response.)

DR. MADAN: Dr. Pazdur, you're --

DR. SINGH: Dr. Pazdur, you're muted for us.

DR. PAZDUR: Thank you. I just wanted to address that comment because I think there's a bigger issue here that has to play out here, from the clinical trial community that this brings up, and that is education of patients and education of investigators, and that's why we highlighted that. And although that might seem rather minimal, I think it's very important that people understand

they shouldn't be going on the trial, and have a careful discussion with the patient, and themselves. If they do not feel that they could take the docetaxel arm if it was allocated to them, they should not be participating in this trial, and that is an end-of-discussion point.

This affects the entire integrity of the clinical trial system if one plays this game of, "Well, I'll go on the drug or I'll remain on the study if I get a certain drug here." We're talking about the integrity of a clinical trial system throughout the world, throughout the United States, and investigators, and patients, and the entire community must take the responsibility of deciding whether they want to go on a trial, and then when they get the results, they have to participate in the trial. It's not, "Well, I'll pick the trial and I'll stay on the trial if I get the arm that I want to go on," and I think that this is an extremely important issue for the committee to discuss.

DR. MADAN: Okay. Thank you, Dr. Pazdur.

Dr. Kraus, we'll get your question in, and then we'll head to our lunch break.

DR. KRAUS: Perfect. Thank you. Can you hear me?

DR. MADAN: Yes, very clearly.

DR. KRAUS: Oh, good. I think Dr. Pazdur led into it a little bit. This situation is a fair bit complicated because of IV versus oral, and therefore, open label, which is unavoidable in this case. So we have to struggle through with a lot of the situation that I'm sure FDA and the sponsor doesn't like, and having been involved with these things, it's very difficult.

One of the key aspects that we're talking about, and I think there's a difference, is the interpretation of optimism, pessimism, and the sensitivity and tipping point analyses, et cetera, and how to look at that. And the question I have can be a larger discussion, and probably will be later. But the question I really have, and we heard the sponsor -- and this would be to the sponsor, but I'm sure FDA will want to comment, and

I'd be interested in their viewpoint as well -- is, what do we have to learn about historic docetaxel data -- there's a fair bit, and the sponsor mentioned it -- around how we should look at these sensitivity analyses, and and how tough -- you can always come up with assumptions that will make a trial look like it didn't work when it did because you change a bunch of assumptions.

I've been through it, and the tipping points are that sort of thing, too. But can we learn something from the historic docetaxel control arm data and look at the control arm in this trial, and say, how in line, out of line? Is there anything we can learn with historical data to know, for guideposts, how the ODAC members here should be looking at this in terms of is it in line, is it out of line, how hard should it be pushed, is it unexpected? That's just a question to the sponsor and FDA.

DR. FRIBERG: Thank you for the question. I'm going to ask Dr. Mehta to directly reply.

DR. MEHTA: Thank you. We looked at a

number of non-small cell lung cancer trials with docetaxel as the comparator arm in open-label situations. Slide 2, please.

On this slide you see CodeBreaK 200 rates of withdrawal prior to and after study drug start, and the right three columns show the data from other trials of pembrolizumab, nivolumab, and avelumab. As you will note, the rates on docetaxel dropout, even prior to study drug start, on CodeBreaK 200 were relatively comparable. Of course, these trials were conducted during different times. CodeBreaK 200 was conducted during the peak of COVID before vaccines were widely available, but generally in that context, these withdrawal rates are consistent across other non-small cell lung cancer trials.

To the broader question of how reliable are these outcomes, I would go back to slide 1, which was shown in Dr. Friberg's presentation that at the end of the day, with all of these challenges, the data from CodeBreaK 200 are incredibly consistent with data from other trials, not only of sotorasib

but of docetaxel. And these docetaxel trials in the right columns, CheckMate 057 or REVEL, or a very recent study, CONTACT-01, the PFS outcomes are remarkably consistent. Thank you.

DR. MADAN: Thank you for that response from the sponsor.

The FDA has a chance to respond now.

DR. SINGH: Thank you. I'd like to make two brief points, Rick, if that's ok.

DR. PAZDUR: Go ahead.

DR. SINGH: Number one, in terms of the historical response rates of docetaxel, they are, in fact, historical, and our assumptions must change over time as data evolves. It is possible that the patients in the docetaxel arm of CodeBreaK 200 overperformed; however, we actually do not have -- and even the slide which the applicant just showed is not technically comparable, some of the trials, because in CodeBreaK 200, all patients had received prior immunotherapy and platinum-doublet chemotherapy, which even the sponsor has considered that this may

have actually augmented the patient's responses to docetaxel.

So the historical knowledge of docetaxel to inform the assumptions of this trial were just that, historical and perhaps non-contemporary. I'd like Dr. Pallavi Mishra-Kalyani to respond, and then Dr. Pazdur, before the sponsor is able to respond because we do have very valid points here, and we should be allowed to complete all of our points before the sponsor responds. Thank you.

DR. MADAN: FDA, and if the sponsor wants to reply afterwards, they can do so.

DR. MISHRA-KALYANI: Great. Thank you, Dr. Singh. This is Pallavi Mishra-Kalyani, FDA, statistics. Certainly, I think the point has been brought up several times that there are differences in the sensitivity analyses conducted by the applicant, as well as conducted by FDA. I think what's most important to understand and to remember in these analyses, really, is why we're doing them, which is, it's that we saw several signs of potential bias and issues with the assessment of

PFS, and we needed to explore these further. We considered which assumptions would be reasonable to make when doing these analyses, but the reason we're most concerned about them is the fact that the PFS benefit that was estimated from the data was marginal. It was incremental to docetaxel, which has already been described as a drug that has marginal benefit to begin with.

So yes, there are differences and, yes, we can talk about the differences in the assumptions. I don't think that optimistic or pessimistic is a valid way to describe these assumptions. I think we have to consider whether or not the data support these assumptions, and FDA has shown that the OS results do support the assumptions that we've made. Lastly, I will just say that no sensitivity analysis can truly mitigate the impact of informative censoring, which is what we've observed in this study.

DR. MADAN: I believe Dr. Pazdur wanted to have a word.

DR. PAZDUR: I'm the only person here that

knows what went on with the original approval of docetaxel and lung cancer, in the agency. And there was a great deal of consternation about this, but the reviewers knew that it had a survival advantage. It had a survival advantage, end of discussion, and we approved it on that basis.

I think we have to get away from this issue, and all of these discussions that we're having here could have been very well mitigated if we really chose the right endpoint here, and that was overall survival in this setting. We wouldn't have to be discussing all of this, and this was pointed out clearly in the FDA slides, that this is a potential for the mitigation of bias, so to speak.

We wouldn't have to be talking about all of these complexities of bias and different sensitivity analysis if we were dealing with either a superiority in overall survival or a noninferiority in overall survival. And I would hope that the field would have moved forward and that we would be able to show a superiority over a drug that was approved 23 years ago, so to speak,

as a kind of where are we going in the field here, so to speak.

Here again, I think we have to take a look at, really, what the basis of the approval of docetaxel was, and it was on not a PFS endpoint, not on response rate, but on a small, but we thought, clinically meaningful endpoint of overall survival. So I'll just leave it at that from the person that has some historical perspective here at the FDA, take it or leave it.

DR. MADAN: Thank you, Dr. Pazdur.

The sponsor has an opportunity to respond.

DR. FRIBERG: I would just point out that we do believe that sotorasib offers something superior to docetaxel, statistically superior progression-free survival, which has a benefit in and of itself, improved response rates and, of course, patients seem to dislike the therapy less. They have a different side-effect profile, and that alone we believe the data supports.

DR. MADAN: Okay. Thank you very much.

With that, I think we will break for lunch.

We will reconvene at 1:15 p.m. Eastern Time. Panel members, please remember there will be no chatting or discussion on the meeting topics with other panel members during the lunch break, wherever that may be. Additionally, you should plan to reconvene for the panel at 1:05 to ensure everyone's reconnected and we can reconvene again at 1:15. Thank you.

(Whereupon, at 12:42 p.m., a lunch recess was taken, and meeting resumed at 1:15 p.m.)

A F T E R N O O N S E S S I O N

(1:15 p.m.)

Open Public Hearing

DR. MADAN: We will now begin the open public hearing session.

Both the FDA and the public believe in a transparent process for information gathering and decision making. To ensure such transparency at the open public hearing session of the advisory committee meeting, FDA believes that it is important to understand the context of an individual's presentation.

For this reason, FDA encourages you, the open public hearing speaker, at the beginning of your written or oral statement to advise the committee of any financial relationship that you may have with the applicant. For example, this financial information may include the applicant's payment for your travel, lodging, or other expenses in connection with your participation in this meeting.

Likewise, FDA encourages you, at the

beginning of your statement, to advise the committee if you do not have any such financial relationships. If you choose not to address this issue of financial relationships at the beginning of your statement, it will not preclude you from speaking.

The FDA and this committee place great importance in the open public hearing process. The insights and comments provided can help the agency and this committee in their consideration of the issues before them. That said, in many instances and for many topics, there will be a variety of opinions. Our goal for today is for the open public hearing to be conducted in a fair and open way, where every participant is listened to carefully and treated with dignity, courtesy, and respect. Therefore, please only speak when recognized by the chairperson. Thank you for your cooperation.

Speaker number 1, please unmute and turn on your webcam. Will speaker number 1 begin and introduce yourself? Please state your name and any

organization you are representing for the record. You have five minutes. Thank you.

MR. MOSBY: Thank you, Mr. Chairman, and members of the committee. On behalf of the Health, Education, Advocacy, and Learning, or HEAL collaborative nonprofit, I am Howard Mosby. I'm a board member and treasurer and will be providing the following oral comments regarding this FDA application for Lumakras, submitted by Amgen. Amgen has been a sponsor for some of our programs that we've had, our educational programs that we've had in the community, but for this particular engagement, I am not being compensated by Amgen for my comments here today.

Now, as you are aware, personalized medicine has become increasingly important in cancer treatment. Targeted therapies which aim to attack specific molecular abnormalities driving cancer growth have shown promise in improving outcomes. African American patients, like all cancer patients, can benefit from these therapies when their tumor's genetic profile matches the available

treatments.

For instance, lung cancer is a significant health problem among African Americans because of the higher rates and incidence in mortality compared to other racial and ethnic ethnic groups in the United States. In Georgia, where our organization is based, the incidence and prevalence of lung cancer among African Americans are higher than the national average, with smoking as the leading cause of lung cancer, accounting for 85 percent of all cases. African Americans have a higher rate of smoking compared to other ethnic groups in the United States, thus the disproportionate incidence of lung cancer in the African American community.

In addition, African Americans may be more susceptible to lung cancer due to genetic factors that increase our risk. African Americans may be less likely to receive family and appropriate health care, including lung cancer screening, diagnosis, and treatment, which can result in higher rates of advanced stage lung cancer and poor

outcomes. Treatment for lung cancer often involves a combination of therapies, which may include surgery, radiation therapy, targeted therapies, immunotherapy, and chemotherapy. The choice of treatment is typically made by a team of healthcare professionals based on individual characteristics.

Innovation is a game-changer in these underserved communities. Our organization has seen individuals that have been misdiagnosed, young persons under 40, and individuals that meet screening criteria that don't get screened until their symptoms reach the worse stages of the disease. And the one thing that jumps out like a sore thumb in this process is that when those individuals receive the state-of-the-art treatment modalities and innovative therapies, we do see survival and success rates to improve their quality of life.

We can state emphatically that survivorship care plans that include new innovative and treatment advances like Lumakras brings positive outcomes and real hope to this population to

survive these deadly diseases that disproportionately affect this community, and we ask that you grant this approval for this new drug application by Amgen. Thank you very much, Mr. Chairman, for allowing me to have these comments.

DR. MADAN: Thank you.

Speaker number 2, please unmute and turn on your webcam. Will speaker number 2 begin and introduce yourself? Please state your name and any organization you're representing for the record. You have five minutes.

MS. CONNERAN: Thank you. My name is Terry Conneran, and I'm with KRAS Kickers. I'm a lung cancer patient that has a KRAS biomarker, and as far as a relationship with Amgen, I have done some consulting work for them as an individual, and KRAS Kickers has received sponsorship from them for a number of different programs, along with a lot of other sponsorships.

First of all, I would like to very much thank the FDA for allowing us as patients, as the public, to lend a voice to this transparent process

because we truly are the people that stand to gain, or lose, potentially, the most in this. We're out here striving to survive. This is an important part of the process, and I appreciate you very much allowing me to be here today.

As I mentioned, I'm a lung cancer patient. When I realized the cancer's bigger than me and we have this commonality, I started an organization called KRAS Kickers literally to bring together patients so we can become empowered about our treatment options and our treatment decisions. That means, literally, the shared decision making is an opportunity for us as far as becoming involved and engaged with a clinical trial.

We so much believe in this that we took KRAS and turned it into an acronym to represent the empowerment that we feel that we need when it comes to living with this disease. We use it as knowledge, plus research, plus efficacy, equals survivorship. Notice I didn't say "cure." We're all out here trying to survive, so that's why I'm here today, is to be able to lend voice to myself

and on behalf of the different patients that are within our group. I myself would not qualify for this particular treatment, so this does not affect me individually; however, over the course of the past 4 years, where we've engaged on a global basis as this group is we've had a number of different people that have been on a number of different modalities of treatments, including these different types of clinical trials.

My understanding of the view of the clinical trial is that there is some concern as far as the biases crossover. As a patient, this is very important to us to be able to have that sense of empowerment that we can cross over or cross out of a clinical trial. I can tell you on behalf of myself, or anybody else, if you were diagnosed with something, and you're put into a randomized situation where you find out that you're just going to get standard of care, not the latest and greatest, wouldn't you wonder? Wouldn't you think? Wouldn't you consider and just, in fact, back out, and potentially push off making that decision to

begin with that chemo treatment right off the front? That would be presumably -- because taking a pill -- or it doesn't matter how many pills over the course of once or twice a day, even 4 times a day -- is a whole lot easier than showing up and being in a chemo chair every 3 weeks and completely losing your life for half of that time.

So I encourage you very much to very closely and very critically review the precedent that you may consider setting, and reviewing this opportunity here, this drug. As you're reviewing it, please give a close eye to the opportunities that may potentially become shut down in the future because it is all about us patients being able to have different opportunities to get involved in clinical trials. And if we lose that flexibility of being able to cross out of it, we're going to be less inclined to do it, and that's going to become limiting for the future treatments for all of us. And that's really all it is we're looking to do, is join together to kick cancers, KRAS. Thank you for having me here today.

DR. MADAN: Thank you for those comments.

Speaker number 3, please unmute and turn on your webcam. Will speaker number 3 begin and introduce yourself?

MS. DONALDSON: Yes. Hello. Can you hear me?

DR. MADAN: Yes. Please state your name and any organization you are representing for the record. You will have [indiscernible] minutes. Thank you.

MS. DONALDSON: My name is Dusty Donaldson. I'm a lung cancer survivor, patient advocate, and the founder of LiveLung, a 501(c)(3) organization with a mission of advancing lung cancer awareness, early detection, and compassion for people impacted by lung cancer. We host a network of educational patient groups to empower lung cancer patients. I'm not a scientist. I'm here today as a patient advocate, speaking on behalf of lung cancer patients. Thank you for the opportunity to speak with you today.

Most people are surprised to discover that

lung cancer is the number one cancer killer, as was I when I was first diagnosed. Lung cancer claims about as many lives as breast, prostate, and colorectal cancers combined. More than 350 people will die each day from lung cancer. After decades of stagnant survival, the outlook is now more promising for lung cancer patients, due in large part to targeted therapies for patients with certain biomarkers such as RET, EGFR, ROS1, ALK, and others.

In 2022, overall cancer deaths were reduced significantly, and according to the American Cancer Society, that change was driven in large part by lung cancer targeted therapies. Specifically, the American Cancer Society attributed the overall survival progress to early detection and treatment advances in lung cancer. Those of us in the trenches with lung cancer patients know that while lung cancer screening protocols exist, 94 percent of eligible candidates are not being screened for lung cancer. When looking at the lung cancer therapeutics landscape, I am persuaded that it is

the targeted therapies behind the improved lung cancer survival rates.

When a newly diagnosed patient first joins our group, they're understandably confused and scared. The seasoned patients rally around them and they share their own stories about how biomarker testing and targeted therapies were a real game-changer for their treatment journey. In that moment, there is an incredibly powerful infusion of hope for that patient. After a patient discovers their biomarker, they connect with other patients in that biomarker community. They are transformed from being confused and frightened to being knowledgeable and empowered. That spark of hope gives them the courage to take the next step, and then the next, in their cancer journey.

Now, that patient may or may not have an actionable biomarker. We understand and accept that sometimes traditional chemotherapy is the only option but, to me, as a patient advocate, chemotherapy is like carpet bombing, whereas targeted therapy is more strategic with less

collateral damage. Targeted therapies are the future, and we know that more biomarkers are being discovered and targeted therapies are being developed that will improve patients quality of life and hopefully lengthen their days here on earth.

Importantly, fear of adverse side effects from cancer treatment, particularly chemotherapy, is one of, if not the, primary reason patients refuse cancer treatment. If given a choice, of course, patients prefer to take pills at home rather than going to the cancer center to receive chemotherapy. The KRAS biomarker is the most common cancer biomarker. It is found not only in lung cancer but in colorectal, pancreatic, and several other cancers. Again, as a lung cancer patient advocate, I encourage and support advancement of targeted therapies, specifically Lumakras, based on the CodeBreaK 200 study, demonstrating that this anti-cancer treatment is less toxic to patients. Thank you so much for allowing me to share my perspective with you.

DR. MADAN: Thank you very much for those comments.

Speaker number 4, please unmute and turn on your webcam. Will speaker number 4 begin and introduce yourself? Please state your name and any organization you're representing for the record. You will have five minutes. Thank you.

MS. WEIR: Thank you very much, Chairman. My name is Debbie Weir, CEO of the Cancer Support Community, an international nonprofit organization that provides support, education, and hope to those affected by cancer. Thank you so much today for the opportunity to speak about this important issue.

On behalf of cancer patients, survivors, and the caregivers we serve, the Cancer Support Community would like to thank you for the opportunity to provide comments regarding the recommendation to update the accelerated approval of Lumakras to full approval. As the largest provider of social and emotional support services for people impacted by cancer, CSC has a unique

understanding of the cancer patient experience. In addition to our direct services, our research and policy institutes are our industry leaders in advancing evidence-based and promoting patient-centered public policies.

We serve all types of cancer patients and their loved ones, including those with lung cancer, the leading cause of cancer-related deaths in the U.S. and worldwide. Given the high prevalence of lung cancer and the scarcity of treatments for locally advanced and metastatic non-small cell lung cancer, harboring the KRAS G12C mutation and the poor 5-year overall survivor rate for metastatic lung cancer and access to drugs that treat this subtype of lung cancer is important to patients and their loved ones.

Having innovative, safe and effective treatment options available would offer additional avenues of consideration, with the ultimate treatment decision always being made between the patient, caregivers, and their healthcare team. While Cancer Support Community does not endorse any

specific product, we do encourage, when appropriate, expanding opportunities that give credence to patients' options and priorities, specifically, the value patients place on both the physical and the psychosocial aspects of their lives.

We appreciate all that FDA has been doing to strengthen this patient-focused drug development program. It is critical that the development of safe and effective therapy options for specific cancer subtypes, which previously had no treatment options, be recognized and elevated as an integral part of the the PFDD program. We ask that the FDA clearly include differences in patient-reported outcomes and side-effect profiles as clinically meaningful and relevant to your approval process.

Even when two drugs have the same efficacy, having the option to choose a different side-effect profile can be extremely meaningful to patients, and also having the choice between oral therapy and IV therapy can be a quality-of-life game-changer. When you think about the impact that regular

infusion appointments have on people living with cancer and their loved ones versus how much easier it is to take an oral medication, it can mean the difference between being able to and not being able to do your activities of daily living. Access to oral therapies can be a health equity issue for the sizable minority of patients. Seventeen percent of cancer patients in our cancer experience registry are very concerned about transportation to treatments and appointments.

We know the patient experience is much broader than patient assessment of disease symptoms, treatment, side effects, and physical functioning. Patient experience also includes psychosocial impacts. We encourage all sponsors to heighten the importance of collecting patient experience data, both preapproval and during postmarket surveillance, by consistently identifying, collecting, measuring, and considering the full breadth of patient experience data to better understand what is really meaningful to patients and their caregivers. We also encourage

sponsors of drugs that are requesting that their accelerated approvals be updated to traditional approvals continue to monitor patients and postmarketing studies to include the build the body, and continue the build the body data on the patient experience.

We would argue that improved patient experience, when observed in a drug that is making accelerated approval criteria, should be considered as a critical part of the subsequent FDA decision-making process. The goal should be to provide meaningful feedback from patients in real time about the issues that may not be identified during the current measures.

We have learned so much from those we serve and support. People living with cancer often feel stigmatized, alone, and overwhelmed with grief and stress. Our oncology psychosocial researchers and others have shown enhancing patients' sense of control can positively impact their psychological well-being. When people living with cancer have more control over the best treatment options for

them, they feel stronger and more hopeful.

Today, we ask that you carefully consider the challenges of those facing KRAS G12C positive NSCLC and the need for a wider array of treatment options for patients. We urge you to support improving access to a broad range of treatment options that would encourage patients to be informed, empowered, and optimistic about their treatment. Thank you so much for your time today, Mr. Chairman. Thank you.

DR. MADAN: Thank you for those comments.

Speaker number 5, please unmute and turn on your webcam. Will speaker number 5 begin and introduce yourself? Please state your name and any organization you are representing for the record. You will have eight minutes. Thank you.

MR. BARANSKI: Hi. My name is Jim Baranski. I'm the executive director of Lung Cancer Foundation of America. We do receive support from industry, and Amgen is one of our supporters. I am not being compensated by Amgen today.

At Lung Cancer Foundation of America, our

mission is principally focused on funding research, specifically young investigator grants, but it's hard to fund research if people don't know that lung cancer is the leading cause of cancer death, so public awareness and patient education are the other elements of our mission. One of the programs that we have is actually built on the shoulders of patients, and these are patients who have the courage to stand up and advocate for lung cancer, both lung cancer awareness, lung cancer research, and just generally living with lung cancer.

In working in that program, I've heard the many, many stories of living with lung cancer and what that means to patients. Words that were once commonplace prior to a lung cancer diagnosis -- words like "scans," words like "progression," words like "toxicity," -- take on a totally different meaning post-diagnosis. I'll never forget the first time that I heard a patient living in Chicago, within miles of a couple of comprehensive cancer centers, share her experience of how the simple matter of IV treatment was, even

though miles away, three bus stops and a couple of hours, and this is in Chicago where it does get cold. So the point is, we've heard about toxicity and we've heard about the difference between take a pill and IV infusion; well, there are layers and layers to that difference, and that difference is a meaningful difference for those living with lung cancer.

The other thing that we hear from patients is, time and time again, how patients are failing to respond to treatment. Patients are failing trials. Patients don't fail responding to treatment, treatments fail in responding to patients. So when we have the opportunity to have a treatment that works for patients at a lesser toxicity, patients welcome that opportunity with open arms. And just a side note, the equipoise discussion, that probably really points to how clinical trial protocols going forward will have to recognize the impact of patients actively involved in sharing and spreading their knowledge on social media, and actually on all platforms of media these

days. So it's no wonder patients want to crossover once they hear the news of another option of less toxicity.

Thank you, Chairman, for your time today, and thank the committee for their time, and thank you to the FDA for making certain that patients that are being treated are being treated favorably.

DR. MADAN: You're welcome, and thank you for those comments.

Speaker number 6, please unmute and turn on your webcam. Will speaker number 6, please begin and introduce yourself? Please state your name and any organization you may be representing. You will have 10 minutes.

MS. ECCLESTON: Hi.

DR. MADAN: We can hear you.

MS. ECCLESTON: Okay. I'm sorry.

DR. MADAN: You're good. Go ahead.

MS. ECCLESTON: My name is Sherri Eccleston, and I'm a 58-year-old cancer patient. I'm not being paid by Amgen or anybody else. I'm not part of any other --

DR. MADAN: Our webcam is not -- okay. Great. There you go.

MS. ECCLESTON: Is it on now? Okay. Sorry.

DR. MADAN: No apology necessary. Go ahead.

MS. ECCLESTON: I'm not part of any organization and I'm not being paid by anyone. I'm here to tell my personal experience.

At age 30, 33, and 35, I had papillary thyroid cancer. I was treated and I was cancer-free from year 2000 until August of 2021, when I had an accident, and I fell down my front stairs, went to the hospital, and they found cancer in my upper-left lung. At that time, I went and I had a lobectomy. Everything was clean. I was stage 1, until I went to the hospital Labor Day weekend of 2022. At that point, I was having different pains in a different part of my body. That day I found out the cancer was back in multiple locations.

From September to October, I had various scans and tests done, and my tumor was sent out for molecular sequencing. It came back the KRAS gene.

My doctor was quite excited when this happened, and he actually called me at about 9:30 at night to tell me about this. And that was a drug he had heard of called Lumakras, but of course Lumakras isn't fully approved, so the insurance company would not approve of me taking it at first.

I had some radiation in November, and then I started carboplatin, permextred, Keytruda, and Avastin. I was very ill. I had pleural effusion, dehydration, I spent time in the hospital in October, half of November, half of December, almost all January, and part of February. I required multiple transfusions, fluids, potassium, magnesium. My blood pressure was up and down. My sugar was out of control, and I had to be put on insulin. I was finally approved for Lumakras in March of 2023.

My last visit before I started Lumakras, I was in bed 24-7. I was only able to make it about 10 feet from my bed to the bathroom. I couldn't make myself a simple sandwich or pour myself a cup of water. I didn't leave the house, except for

going to the doctor or the hospital. When I did go to the doctor, my daughter had to drive into my yard to the bottom of the six steps in front of my doorway, and I painfully, slowly made my way down those six steps, to the car door that was right at the bottom. And when I got to the doctor's office, I was immediately put into a wheelchair. That was my quality of life, nothing but bed and going to the doctor.

I started Lumakras, and my follow-up appointment 3 weeks from there, I walked down the front steps to my driveway, got into the car, walked into the doctor's office, and after the doctor, my daughter and I actually went to the diner. That was my first outing in all those months.

I still suffer from effects of neuropathy and have issues with other things that I have to take care of, but while I know you were trying to make sure these studies were done right, my doctors, pre- Lumakras, did not think I would be here today, and neither did I. Any of my friends

that came to visit me, have since told me, when they left the house, they sat in their car crying before they could even pull away from the house because they were afraid it was going to be the last time they saw me.

All of my tumors have since shrunk. I had a scan in July of 2023, and it said, "near complete resolution of disease." Every time I see any of my doctors, or nurses, now they are completely shocked over and over again at how well I look. This is why I feel compelled to speak to you today, to ask -- no, plead -- for approval of Lumakras. Without Lumakras, I am sure I would not be here today. Thank you so much for your time, Mr. Chairman and the committee. Please approve Lumakras.

Clarifying Questions (continued)

DR. MADAN: Thank you for sharing your story.

The open public hearing portion of our meeting has now concluded, we've had all six speakers, and we'll no longer take comments from

the audience. We have about a half hour of time here, and what we'll do is reopen the floor for any clarifying questions or discussion further from the committee, if appropriate and if we have time.

As we have additional time, we will now take these remaining clarifying questions, if there are any. Again, please use the raise-hand icon to indicate if you have a question, and remember to put your hand down after speaking. Please remember to state your name for the record before you speak and direct your questions specifically to a presenter, if you can. If you wish to have a specific slide displayed, please let us know the slide number, if possible. And as a gentle reminder, it would be helpful to acknowledge the end of your questions with a thank you, and at the end of your follow-up questions, if you have any, "This is all for my questions."

We can move on to this portion if we have -- let me just see what happened here. I think my Zoom screen went blank. Hold on a sec. It's always something exciting.

(Pause.)

DR. MADAN: I am not seeing my screen here, so I apologize for the technical issues. I will clarify if anyone has any questions.

DR. SINGH: Dr. Madan, this is Harpreet Singh from the FDA. I do not see any hands raised. Oh, I do. I apologize. I'm starting to see hands raised.

Do you see them now or would you like my assistance?

DR. MADAN: Yes --

DR. SINGH: I can tell you -- do you see them?

DR. MADAN: Thank you.

DR. SINGH: Do you see them or would you like for me --

DR. MADAN: You can tell me who the first person is.

DR. SINGH: First, it appears to be Dr. Gulley, followed by Dr. Spratt, followed by Dr. Vasan, followed by Dr. Shaw, followed by Dr. Nieva, and I can put that to you in the chat on

the backend. It appears Dr. Gulley has his hand raised.

DR. MADAN: We'll go ahead, and I'll sort out my technical issues.

Dr. Gulley, please proceed with your question.

DR. GULLEY: Thank you. James Gulley, NCI. I just wanted to come back to the question that I asked earlier to see if the applicant had a chance to get the data, specifically on the early crossover, if we can have clarification on the number of patients that crossed over that did not progress -- I believe it was 19 -- on the initial BICR evaluation, and what the RECIST responses were for the final BICR evaluation for those, and if any of those have a progressive disease on that final evaluation.

DR. FRIBERG: Yes. Thank you for the reminder. We're going to need about 15 more minutes, but we should absolutely be able to have that for you shortly.

DR. GULLEY: No problem. Thank you.

DR. MADAN: Alright. Thank you very much, for the sponsor. We will make a note to come back to that at the end and allot time.

I am back. I wasn't gone, but my screen was, I guess, so I apologize for that.

Our next speaker with a question is Dr. Spratt.

DR. SPRATT: Thank you, two very simple questions. The first is for the sponsor. Was any quality-of-life data beyond progression collected, even if only a subset?

DR. FRIBERG: I'm going to ask Dr. Stollenwerk to comment on your question.

DR. STOLLENWERK: Hello. My name is Bjorn Stollenwerk. I'm a director of health ec, and I work with Amgen. Most of the quality-of-life data was not measured beyond progression. There was only one single exception, and that was the EQ-5D data, which was also in long-term follow-up.

DR. SPRATT: Thank you very much. Did you present -- and I apologize if I missed it -- the data after progression, even with EQ-5D?

DR. STOLLENWERK: It was not a trial endpoint, so those data were measured for different purposes. We don't have a long-term presentation ready, I think, to present here.

DR. MADAN: That's the reply there.

DR. SPRATT: Thank you.

DR. MADAN: Dr. Spratt, thank you for your questions.

Our next question is Dr. Vasan.

DR. VASAN: Hi. Neil Vasan, Columbia University. I just wanted a little more granularity on a point that Dr. Madan brought up earlier, and this is a question for the FDA.

In the briefing document, it says on page 16 that FDA found the proposed study design generally acceptable but expressed concerns that the targeted 3.2-month difference in median PFS would not be considered clinically meaningful. So the applicant had said, in response to Dr. Madan's question, that that 3.2-month benchmark was before the CodeBreaK 100 results, and that that number was sort of going to be more tempered.

But I guess the initial FDA assessment of that 3.2-month benchmark, does that statement need any qualification or was that number sort of deemed as an absolute, like this is what is clinically meaningful, regardless of the results of CodeBreaK 100?

DR. SINGH: Okay. Thank you. Harpreet Singh, FDA, and I'll respond to this on behalf of the FDA. We do not have a definition of what we consider a clinically meaningful PFS. What was left out of that discussion, really from both sides, is that in a refractory setting, we typically do ask for an overall survival endpoint, particularly in patients with poor prognosis, patients with unknown prognosis, as in the case for patients with KRAS G12C mutations.

This was a head-to-head design, so we did feel that the target in median improvement of PFS may have been clinically meaningful if they, in fact, reached that benchmark, but we always assess the totality, and particularly the overall survival. So we always qualify our statements if

an applicant or sponsor chooses to move forward with a PFS primary. So we did feel that that 3.2 months could be considered clinically meaningful as a median, and, of course, as we discuss in our presentation, we look at other measures of a PFS effect such as hazard ratios, such as medians. But what's left out of that conversation is, basically, we a priori ask for survival as the primary, as we did with this sponsor, as we do with refractory trials.

DR. VASAN: Thank you.

DR. MADAN: Thank you.

I'll give a chance for the sponsor to reply if they have anything to say.

DR. FRIBERG: I would only add that the initial study design was powered to look at overall survival. After the CodeBreaK 100 data became available, there was strong feedback from not just the investigators on the study, but regulators around the world that crossover was something that should be implemented for patients. And the implications, of course, of that were that we would

also reduce the study and have PFS be the primary endpoint.

DR. MADAN: Okay. Thank you.

Dr. Shaw, do you have a question?

DR. SHAW: Yes. I think this is really a question for the FDA and just making sure I understand the wording of the question and how I should answer it.

DR. MADAN: Just as a point of order, we will have an opportunity to clarify the question before the voting.

DR. SHAW: Oh, okay. So maybe I should hold it then.

DR. MADAN: We can come back [indiscernible], unless you have a question specifically more for discussion purposes.

DR. SHAW: I see. Thank you for clarifying. So just for the record, this is Pamela Shaw, and I'll hold my question for the proper time. Thank you.

DR. MADAN: Thank you very much.

Dr. Nieva, are you there?

DR. NIEVA: Yes. Thank you very much. I just want to give another opportunity for the applicant to clarify things about the second analysis of the blinded independent central review. There were statements made from the FDA that there was a lack of clarity regarding exactly what happened around this time, what actually triggered the second analysis.

It does appear that there was an interim analysis performed that Amgen was privy to before deciding on engaging in a re-analysis. So I just want to confirm that that interim analysis was specified in the protocol that there would be an interim analysis, and I just want to clarify that that was actually in the imaging charter for the protocol. If there are things that the applicant would like to say now to make things seem more transparent as to exactly what happened around that time, it'd be appreciated. Thank you.

DR. FRIBERG: Yes. Thanks for the opportunity to clarify. Just to be clear, there was only one interim analysis. It was prespecified

in the protocol, and that was, of course, blinded to Amgen. It went to the DMC. It was independently run by the imaging vendor working with the DMC, and the DMC recommended to continue the study as planned. Around the same time, there was this observation in the aggregate data that there was some discordance, and timelines, and connectedness between raw event rates. This was flagged to the imaging vendor, who went through an independent review process. They called it their reader performance monitoring. That independent process ultimately led to data corrections.

So those data corrections -- there were 11 data points that were part of the interim analysis -- were then corrected, and the interim analysis was re-run with the corrected data. That went to the DMC, and the DMC did not recommend any changes to the plan. That being said, the FDA and Amgen discussed this, and given the potential for the introduction of bias by this initial communication, based on aggregate data to the imaging vendor, it was decided that the right thing

to do would be to do a 100 percent re-read of all the analyses. So all of this discussion is about this interim analysis, where ultimately we followed the data monitoring committee. The final analysis is based on a 100 percent re-read of all the scans.

DR. MADAN: Okay. Thank you for that reply.

Dr. Nieva, any follow-up questions?

DR. NIEVA: No. Thank you. I think I understand.

DR. MADAN: Alright. I'd like to ask a question now, if that's ok. Ravi Madan, National Cancer Institute. There have been some kind of allusions to this, but I don't think we've explicitly talked about the size of this study. There was a benchmark analysis at one year that showed a strong trend favoring the experimental intervention, but there was only 37 patients to evaluate at that one-year mark, and that was kind of consistent if you went beyond, I believe, 7 months. I guess the statisticians on both sides from the FDA and the sponsor, if they could talk about how size factors into this process of

analysis here and how it makes this discussion more complicated with some of these potential biases and dropouts. Thank you. For ease, we'll start with the FDA, and then the sponsor can go last.

DR. SINGH: Okay. I would like to ask one of our senior biostatistical colleagues to join us on video to respond regarding sample size. I see Dr. Amatya is here. Thank you.

DR. AMATYA: It's Anup Amatya, FDA. The CodeBreaK 200, the result that we're discussing, it had adequate power to analyze PFS as designed, which was when you revised, it was targeted, the magnitude went down to 3.2 months; however, with the revision, the sample size was reduced, and the power for OS analysis was about 50 percent or 58 percent. So from that perspective, sample size was not adequate for OS analysis; however, the sample size was adequate for PFS.

So interpretation regarding PFS from a sample size perspective is not an issue, but the issue is the dropout after the trial has been started. So there was a significant amount of

dropout midway through the study. When I say dropout, it's either because patients withdrew or were censored because the progression was called by investigator before the BICR. So for the primary endpoint analysis, the censoring issue was significant after 6 or 7 months of follow-up. That created a lot of uncertainty regarding the interpretation.

DR. SINGH: Dr. Amatya, to clarify, are you saying that from a statistical perspective, the further ends of the curve, after a long-term follow-up, are less reliably interpreted because of the very small numbers of patients remaining on each arm? Is that accurate?

DR. AMATYA: Yes. In particular, this is even more so for a docetaxel arm, where only 7 patients were left or comparable.

DR. SINGH: Okay. And the question around sample size, Dr. Madan, are you asking if the sample size had not been reduced, if that reliability would have been less uncertain, or --

DR. MADAN: No. I think my untrained

statistician perspective is that there's a lot of turbulence with the dropout and the censoring, and had this been a larger trial, perhaps the data would be more convincing.

DR. SINGH: Yes. I'm glad you raised this question, and this is something that the FDA worked through, obviously, before we chose to bring this to a committee and throughout our review process. I think that the truth of the matter is we will never actually know; however, there's no reason to believe that more patients in a larger sample size would have impacted the trends that we're seeing very early on, the high rates of early dropout, which, again, were mitigated ultimately, and the majority by the institution of crossover, which the FDA worked with the sponsor to institute.

What I want to add to shade this conversation is many of these discussions we simply would not be having if the effect size of the drug in question, sotorasib, was greater in magnitude. Even though we're in a head-to-head setting, we do have a marginal comparator here, as everybody, both

sides have acknowledged. Thank you.

DR. MADAN: Thank you for the FDA's response.

Does the sponsor have anything they'd like to reply to that question?

DR. FRIBERG: Thank you for the opportunity. I'd like to ask Dr. Koch to comment on this, particularly with regard to that 12-month time point where there was a 25 percent rate of progression free and alive versus 10 percent for docetaxel.

DR. KOCH: Gary Koch, biostatistics department, University of North Carolina. Can you bring up the Kaplan-Meier curves for the comparison of PFS that were in the main presentation? What I was going to try to clarify is that at the 12-month milestone, there may only have been 37 patients remaining at risk, but you can see that many of the patients in the docetaxel group, an estimated 90 percent, slide 2, basically already had PFS events. So a major reason for the decrease in sample size is previous PFS events, particularly in

the docetaxel arm.

So the estimates at 12 months are actually based on all of the data in terms of how Kaplan-Meier estimates are calculated. Certainly, patients with censoring do not contribute beyond the time of censoring, but these estimates are based on all of the data, particularly the patients that had the PFS events and for the patients with censoring, as long as they were followed. So the 37 there is mainly driven by patients that had previous PFS events.

Then if we want to go over to slide 1 just as additional clarification, even though the difference in medians, as shown in the lower right-hand corner, is only 1.1 month, the difference in 40th percentiles is 2.8 months and the difference in the 25th percentile is the 2.9 month and, again, patients are contributing to these estimates of percentiles.

The median is a horizontal difference at the 0.5 point between the two curves, but that horizontal distance varies a lot as you move down

to the lower quantiles. And as I just said, at the 40th percentile, which is near the median, the difference is 2.9 months, and then, as the sponsor pointed out, at the 12-month milestone, the treatment difference is 14.7 percent months. And you can see that the upper limit on docetaxel at 12 months is somewhat bigger than the lower limit for the sotorasib, and this will tell you that, basically, these estimates have reasonable precise estimation, and if you were to do an informal comparison at 12 months, it would nearly have a p-value below 0.05.

I haven't done that calculation but, again, as the sponsor indicated, at 9 months, the difference is 14.4 percent, and at 15 months, it's 11.2 percent, so it's a relatively similar difference in milestones throughout the range from about 8 months to 14 months.

DR. PAZDUR: I have a question. Were any of these analyses prespecified, the 12 month analysis, landmark analysis? I doubt it. Isn't this akin to shooting an arrow on the wall and then drawing a

target around it?

DR. KOCH: Well, that is why --

DR. PAZDUR: What's the validity of this? You obviously have looked at the data here already, and you're actually conferring -- you gave a p-value here, actually -- statistical significance to a non-prespecified analysis after taking a look at the data. Right?

DR. KOCH: Let me try to clarify that. When you have Kaplan-Meier curves and you have an overall difference with a hazard ratio that achieves a p-value of 0.003, it is indeed of interest to identify what parts of the Kaplan-Meier curve are driving that difference. And even though I don't know what the nominal p-value there is, it is still useful to look at confidence intervals at different points in time.

I think that the FDA did indeed do that, although I don't know that they did the difference in Kaplan-Meier curves at different points in time. But the key point is that although the 12-month milestone is arbitrary in some sense, the more

important point is that the vertical distance between the Kaplan-Meier curves is reasonably stable between 8 months and 15 months, whereas the horizontal difference at the median ends up becoming much larger as you move down towards the 40th, or the 30th, or the 20th percentiles. Again, when you have a significant hazard ratio, you are able to interpret differences in Kaplan-Meier curves, whether those comparisons are formal or not, and in this case, they are informal. I certainly agree with that.

DR. PAZDUR: Okay. Fair enough.

DR. MADAN: That's the key point to the question there.

DR. MISHRA-KALYANI: Could I have an opportunity to respond to Dr. Koch's comment regarding the censoring at 12 months?

DR. MADAN: Yes.

DR. MISHRA-KALYANI: I'll make it very brief. If you could please bring up slide 15 from the FDA main presentation?

Dr. Koch's comments were that, at 12 months,

there were 37 total patients left, and the majority of these patients were removed from the risk set due to events. If we look at slide 15 from FDA's main presentation, we see that of the patients that were removed from the risk set from the docetaxel arm, 67 were censored and seven remained.

DR. MADAN: Okay. That's important.

DR. MISHRA-KALYANI: I'm sorry that the slide hasn't come up, but in the docetaxel arm, especially, we see that the rate of censoring is very, very high, and this reduces our confidence in the curves, particularly in the later point. The landmark analysis -- thank you very much for bring up the slide -- certainly is arbitrary, and we do recognize that it's important to look at the full curve, but we also need to understand the reliability of the data in the later half of the curve when making inference from that data. Thank you.

DR. MADAN: I know that I asked you to be brief, but I'm going to prolong because I think these are important points. You mentioned some

numbers. Can you just repeat that again, now that we can all see the --

DR. MISHRA-KALYANI: Absolutely. Sure. If you look at the 12-month time point, we see that there are 7 patients left in the risk set, and we started with 174 in the docetaxel arm. Dr. Koch mentioned that the majority of patients who were removed from the risk set were removed because they had events; however, I think it's very important to acknowledge that 67 patients on the docetaxel arm were removed from the risk set because they were censored, not because they had an event.

On the other hand, in the sotorasib arm, there are 30 patients left at the risk set at 12 months -- that's the information we're using to inform our landmark analyses -- but only 32 patients were censored prior to that. So we have a lot more events happening in the sotorasib arm, informing the landmark analysis, than we have in the docetaxel arm.

DR. MADAN: Okay. Thank you for clarifying that.

Again, I want to give equal time to the sponsor if they have anything they want to say in response to this slide or that comment.

DR. FRIBERG: Thank you so much. If we could bring up slide 2? I just want to comment again, we're mixing different analyses, and I wanted to point out that with regard to the BICR PFS analysis, we accounted in our analyses for these 49 and 73 patients. We've looked at this a variety of different ways, and every technique that we've used has shown roughly the same relative risk reduction of around 30 percent.

So again, I think we can speak about hypotheticals, we can look at post-randomization factors, and we can pick different milestones, but the primary endpoint of the study, when you look at it by Kaplan-Meier and you look at hazard ratios, appears to be robust.

DR. MADAN: Okay. Good. I'm glad you guys both had a chance to address that issue.

Before we get back to Dr. Gulley's question, I just want to make sure that everyone has a chance

on the panel, and Mr. Pantelas has been waiting patiently, but he will have the last question, and then we will go back to Dr. Gulley's question, which I believe the sponsor is working on the response and we appreciate them for doing that.

Mr. Pantelas, if you can ask your question.

MR. PANTELAS: Thank you. Jim Pantelas, patient advocate. In a way, I'm worried that we're throwing the baby out with the bath water in this whole conversation because I wonder if there is a reasonable structure in which to actually compare PFS and OS, and if it also requires us to be unreasonable in our patient expectations. I think in this case we're being unreasonable.

There was a comment after my last question, where someone said, essentially, that patients need to be educated because we can't do this research without their compliance, but patients are complying to a level that's reasonable. It's not reasonable to ask patients to sit in an arm that has a lot more side effects or that requires a lot more of them in an unreasonable fashion. And what

I'm hearing is comparing the results of the trial drug results to historical data on the doxy isn't going to work, and I don't know why. We've worked with that drug for a long time, and I don't think that you're going to educate patients, who are trying to save their lives, to go on a on a trial with something that's not more viable than the trial drug.

DR. MADAN: I'm just trying to distill that into a question, and I think it was a good commentary as well. But I think at the end, your question is, how can we expect patients to go on randomized trials and can we just use historical controls?

MR. PANTELAS: And how can you compare these two things?

DR. MADAN: Yes, that's the tricky part.

MR. PANTELAS: How should this have been designed?

DR. MADAN: Yes. No, that's kind of the tricky part of all of this and clinical research in general.

I will ask the FDA --

DR. SINGH: Thank you.

DR. MADAN: -- and give the sponsor a chance to reply to the the question that I'll pose, which is the reliability of historical controls and how they can be used to interpret this data.

DR. SINGH: Okay. Well, thank you for the question, Mr. Pantelas, and I have really appreciated your comments throughout today's advisory committee, and we share your sentiments regarding both patients and investigators acting in what they believe is their own best interest and trying to access what they believe are life-saving therapies.

I think the question that you're asking really gets around why we must conduct randomized-controlled trials, and you're asking basically why we cannot just rely on the single-arm data, the response rate data, or perhaps you're not asking that. So let me just share with you my -- I'll get through my part, and then please feel free to respond.

But I will say that, first of all, there are no good historical benchmarks for patients treated with docetaxel who have been previously treated with immunotherapy and chemotherapy. They are emerging, but because immunotherapy and chemotherapy became based on landmark trial results, demonstrating overall survival benefits, I will add, in a frontline setting, overall survival -- and you did mention survival -- those therapies are only about 5 years old in terms of first line. So the data that we actually have to compare historical response rates of docetaxel after immunotherapy plus platinum-based chemotherapy are limited.

The reason that randomized trials must be conducted, we believe, is because from a statistical standpoint, you can only interpret time-to-event endpoints best in the setting of a randomized-controlled trial, whether that time-to-event endpoint is progression-free survival or overall survival. What we hope to do is shift this conversation not so much to whether we should

be approving sotorasib -- because, remember, we are not actually asking the committee this because there are many regulatory pathways available to us -- what we are asking is actually to look at the trial itself of CodeBreaK 200.

We do believe this is a very challenging space to conduct clinical trials. We have an embarrassment of riches at times, and there are things that could have been done to mitigate some of what we saw here. One key thing that could have been done is real-time assessment of progression before crossover. Also, if there was belief that crossover should have been instituted from study start, as you saw in a competing trial that is ongoing now, that would be another mitigation strategy to mitigate patient dropout. That trial has a 2-to-1 randomization.

When we engaged with the applicant after we had the final top-line results from CodeBreaK 200, we discussed a variety of methods to maintain equipoise in their ongoing trial, but both the FDA and the applicant, in fact, were blinded, as they

mentioned, to many of the patterns of behavior that were already ongoing, that had already escaped, basically, the confines of this clinical trial and their own strategies to mitigate bias.

We are not looking here to place any blame on patients or investigators. We are saying that in today's information age, yes, things could have been done better, but ultimately because the effect size is quite marginal here, you would have expected more of an effect size. And I don't think it's a universally accepted concept that everybody wants oral therapy versus IV. The toxicities are different, but not each individual patient -- it's not a monolithic experience, and we appreciate that as well.

So we cannot, a priori, decide for patients what is a better option. That is why these clinical trials are conducted. And as Dr. Pazdur so rightly pointed out earlier, we do believe that patients are seeking therapies that prove a superiority over the existing historical single-agent chemotherapies. Thank you.

DR. MADAN: Thank you, Dr. Singh.

I do want to make this balanced, and if the sponsor has anything they want to say in response to the comments made, please go ahead, but we do want to also get to the answer to Dr. Gulley's previous question, which I also know the sponsor's been working on.

DR. FRIBERG: Yes. Thank you very much, and I'll make my response brief. Just as a quick clarification, the other KRAS G12C study that's being referred to, I believe that the addition of crossover in that study was added through an amendment. And again, we're all victims of time and place when we run our studies, so there is the potential that could have benefited from some of the experience that we've gone through in our program.

I would just say that with regard to real-world evidence, we have a a variety idea of real-world evidence sources. Those are quite helpful in putting data into context, but as Dr. Singh nicely pointed out, they're not viewed as

substitutions around the world, currently, for randomized-controlled studies.

If it's alright with you, I'll pass the podium to Dr. Mehta to answer the question directly from before. Is it ok to do that right now?

DR. MADAN: Yes, that would be great, and thank you very much for doing that.

DR. FRIBERG: Thank you.

DR. MEHTA: Thank you. I would like to address Dr. Gulley's question around the target lesion percent changes in the docetaxel arm patients who crossed over early. If the slide cores can please bring up slide 3?

We very rapidly took a look at this, just QC'd [indiscernible], and let me walk you through the slide. These are the 19 patients that were referred to in the FDA briefing document, where crossover occurred prior to a BICR PD call. These lesion sizes are changes from the nadir to the last BICR scan prior to crossover, and here are the percent changes in lesion size.

Dr. Gulley -

DR. MADAN: Dr. Gulley, you have an opportunity to ask a follow-up since this is the data you wanted.

DR. GULLEY: Yes. Thank you so much for this. It looks like there is a displaying of results as one can often see, and it also could be that these were different target lesions than the ones that were used in the COP, so this is very helpful. Thank you so much.

DR. MEHTA: Correct. Yes. These were the target lesions followed by the BICR. Thank you.

Questions to the Committee and Discussion

DR. MADAN: Okay. We appreciate the work that went into pulling that data up during the meeting, and I think with that, we will end our open discussion session or I guess clarified questions session. The committee will now turn its attention to address the task at hand, the careful consideration of the data before the committee, as well as the public comments.

Dr. Joyce Frimpong will address the

instructions for voting.

DR. FRIMPONG: Thank you, Dr. Madan.

This is Joyce Frimpong, designated federal officer. Voting members will use the Zoom platform to submit their vote for this meeting. If you are not a voting member, you'll be moved to a breakout room while we conduct the vote. After the chairperson reads the voting question into the record and all questions and discussions regarding the wording of the vote question are complete, we will announce that voting will begin. A voting window will appear where you can submit your vote. There will be no discussion during the voting session. You should select the button in the window that corresponds to your vote. Please note that once you click the submit button, you will not be able to change your vote.

Once all voting members have selected their vote, I will announce that the vote is closed. Please note there will be a momentary pause as we tally the vote results and return the non-voting members into the meeting room. Next, the vote

results will be displayed on the screen. I will read the vote results from the screen into the record. Thereafter, the chairperson will go down the list and each voting member will state their name and their vote into the record. Voting members should also address any subparts of the voting question, including the rationale for their vote.

Are there any questions about the voting process before we begin?

(No response.)

DR. FRIMPONG: Since there are no questions, I will hand it back to Dr. Madan, and we can begin.

Back to you, Dr. Madan.

DR. MADAN: Okay. Thank you.

Now, I will read the question. It's only one question today for the committee, and this is, again, a voting question. The question is -- and this is the specific question that we're voting on -- can the primary endpoint, progression-free survival for blinded independent central review, or BICR, as we've called throughout the meeting, be

reliably interpreted in CodeBreaK 200?

Do we have anyone who wants to clarify anything about this question? I know that there was a question earlier and I deferred it to now, so please let us know. You can just weigh in. I think it was Dr. -- I apologize. Maybe it's not a remaining question.

DR. SHAW: It was Dr. Shaw.

DR. MADAN: Dr. Shaw.

I think just for emphasis purposes, I'd like to clarify -- and I think I know the answer -- we're not making an approval discussion today, or decision, or voting on potential approval. We're asking a very specific question about this specific trial and this specific data set; is that correct?

DR. VELLANKI: Hi. This is Paz Vellanki from the FDA. Yes, that is correct. We are not asking the committee to opine on whether or not we should convert the accelerated approval to a traditional approval for sotorasib, but really we are interested in hearing whether or not we believe

that the progression-free survival per BICR endpoint can be reliably interpreted, meaning can we say for sure that there is a PFS benefit of sotorasib over docetaxel, and can we quantify that effect.

DR. MADAN: Okay. Thank you for affirming that question.

Are there any other questions from anyone else? And I don't see any hands raised either.

DR. SHAW: I'm sorry to speak without being called on. It's Dr. Shaw. My hand is raised, I thought.

DR. MADAN: I'm sorry. I didn't see it. I apologize. Go ahead, Dr. Shaw.

DR. SHAW: Okay. I just wanted to ask a clarifying question about the question here, and perhaps the person who just spoke, maybe she answered it, but I just want to double check.

When I think about this question of whether or not I can reliably interpret the results of CodeBreaK 200 regarding progression-free survival, I think about how I normally interpret results from

an open-label trial, which is I consider the primary endpoint, but I also am considering the totality of evidence more broadly, particularly in open-label studies. So I wasn't sure if I really am being asked --

DR. PAZDUR: We're not asking an approval question; we're asking a question about this specific endpoint.

DR. SHAW: Right. But the person who just spoke before you said I'm being asked whether or not I think there is a benefit, and I don't know -- whether or not I can interpret it is different than what my conclusion is. So I want to make sure --

DR. PAZDUR: The question is written in vernacular English, so we're looking for the effect on an endpoint here --

DR. SHAW: Okay.

DR. PAZDUR: -- and I think it's clearly stated. We're not asking about an approval, the totality of evidence, et cetera. We're asking about the primary endpoint on this trial. Okay?

DR. SHAW: Whether or not there was a benefit specifically on the primary endpoint --

DR. PAZDUR: Correct, is there an improvement.

DR. SINGH: No. That --

DR. PAZDUR: Not benefit; is there an improvement.

DR. SHAW: Improvement. Excuse me.

DR. SINGH: And can it be reliably quantified.

DR. SHAW: And reliably quantified.

DR. PAZDUR: Right.

DR. SINGH: When we talk about the interpretation of the end point, it's not just is there an effect; it's can you actually interpret the magnitude of effect? Can you reliably interpret? Do you believe the hazard ratio? Do you believe the median? Do you believe that it's robust? And that final slide that we showed with the various measurements of the primary endpoint, do you find that we are able to reliably interpret the data supporting that endpoint, and ultimately

the primary endpoint of the trial, which is progression-free survival by blinded independent central review? That endpoint, and that endpoint alone.

DR. SHAW: I mean, what I'm hearing here, just to make sure I understand, it's a bit of a tall order. An open-label trial and how I interpret that primary endpoint is different than how I interpret a randomized, double-blinded trial. We're all in this context of an open-label --

DR. PAZDUR: It is what it is.

DR. SHAW: Yes, ok. Maybe I'm just stalling at this point, and I should --

DR. PAZDUR: It is what it is. Okay?

(Crosstalk.)

DR. SINGH: I think what Dr. Shaw is speaking to is the magnitude of benefit, which one may expect to see in a head-to-head design in an open-label trial, or the totality versus what you would see in a double-blinded trial, which is exceedingly rare in oncology. You may view the totality of evidence differently; hence, the

survival; hence, the PRO data. We are not asking about that. We are asking solely about the integrity, the fidelity of the primary endpoint, which is the only endpoint which is statistically tested. We do not label descriptive information, typically, that lacks statistical rigor.

So in this case, the entire trial rests on the integrity and the fidelity of the primary endpoint, and we're asking if you believe, based on both sides presented here, you can reliably interpret those findings. Thank you.

DR. MADAN: Dr. Shaw, I'll just add that after the vote is done, everyone will have a chance to characterize their answer, and this is a good chance for you to speak about the nuances of whatever vote you make.

DR. SHAW: Thank you for that advice, and that ends my question. Thank you.

DR. MADAN: Okay. Great.

Are there any other questions? I don't see any other hands raised; so jump out, otherwise we'll move on.

(No response.)

DR. MADAN: So if there are no further questions or comments concerning the wording of the question, we will now begin the voting.

DR. FRIMPONG: We will now move non-voting participants to the breakout room.

(Voting.)

DR. FRIMPONG: Voting has closed and is now complete. The voting results will be displayed.

(Pause.)

DR. FRIMPONG: There are 2 yeses, 10 noes, and zero abstentions.

DR. MADAN: Thank you.

We will now go down the list and have everyone who voted state their name and their vote into the record. You may also use this opportunity to include a rationale for your vote.

DR. FRIMPONG: Dr. Madan, give us a second for the polling to come up on the screen. It should be up momentarily.

We're good now, Dr. Madan.

DR. MADAN: Okay. So again, we'll just go

down the list. Again, state your name, your vote, and please feel free to add any background or rationale.

Doctor Conaway?

DR. CONAWAY: Yes. Mark Conaway, University of Virginia. I voted no. No one expects a perfect RCT, but what we hope for is a small number of issues in trial conduct and an effect large enough to withstand the uncertainties caused by those issues. For this trial, we seem to have the opposite, a large number of issues that cloud the interpretation of a small observed effect, so I voted no.

DR. MADAN: Okay. I guess I'm next. Ravi Madan, National Cancer Institute. I voted no. The question before the committee today is not one of the efficacy of sotorasib in lung cancer, but rather, specifically, the ability to interpret data from a relatively small clinical trial conducted with a highly anticipated agent in a hyper information age where both patients and providers had high expectations.

Given that we had hours of statistical permutations discussed that could change interpretations, I had to vote no on the reliability of the PFS benefit from this study. The factors that contributed to the lack of certainty really come from, again, the small size, investigator conduct, and the small 5-week PFS benefit. I do think if the PFS benefit was much greater, this would have been a much shorter conversation.

But this question will not be limited to this study in the future. Industry and investigators must work together to ensure clinical trials are conducted competently so that we can glean the best data to advise our patients based on outcome data and not presumption. The sponsor is to be commended for choosing the appropriate and active control arm in the study, which is not always the case for highly anticipated drugs in this day and age.

But clinical investigators must comply with the spirit of the protocol and provide necessary

education as part of the informed consent process so that once enrolled, patients have the comfort and confidence to continue with the study. Only then can we move forward with new therapies that have demonstrated convincing clinical benefit without question. Data fidelity must begin with the fidelity of the investigators to the protocol. Thank you.

I'll move down to the list to Dr. Rosko.

DR. ROSKO: Ashley Rosko. I voted no. My vote reflects the stance that the results should be informed by a well-controlled trial. The process, to me, by which the radiologic re-read was performed, and triggered a subsequent reanalysis, impacted the integrity of the study, to me, and it's opened up other questions about that immediate dropout, the crossover without bigger confirmed progression.

This impact and perception of study arm equipoise is really hard to measure post hoc, and I do appreciate the efforts that were in the discussion today regarding guidance from the FDA

and how to be able to better mitigate these strategies, whether for applicants, for investigators, and for all.

I did want to just mention from a study angle that the clinical perspective, I did appreciate that from the applicant and also from the patients in terms of having well-tolerated therapy, and that does provide options for patients. But ultimately as a clinician, I wanted to be very confident in the data that I'm interpreting for patients, and that any therapy will provide a substantially better, speaking to the effect size, or longer life lived for the patient.

DR. MADAN: Thank you, Dr. Rosko.

Dr. Nieva?

DR. NIEVA: I also interpreted the question here very narrowly, and I compliment the statistical teams of both the FDA and the sponsor for the work done. I voted yes because the study met its primary endpoint based on the intent-to-treat analysis, and ultimately we have to

take the statistical plan as it is written and analyze things according to what was planned. I think the post hoc analyses are informative, but they ultimately don't change the benefits that were in fact observed, and I don't think a type 1 error occurred here. Given the corroborating evidence, I have confidence that the drug does have a PFS benefit over the comparator in this case.

I will like to add a [indiscernible] that like Dr. Rosko, I am also concerned with the quality of the blinded independent central review and the substantial variation that occurred between the first and second interpretations. I do think that needs greater scrutiny from the FDA and greater transparency from the applicant, but I accepted the results as presented, though think they should be subject to greater auditing. Thank you.

DR. MADAN: Thank you, Dr. Nieva.

Dr. Shaw?

DR. SHAW: Yes. Pamela Shaw at Kaiser Permanente Washington Health Research Institute. I

voted yes, as well. I voted yes because I think there were a couple of different discussions here today. Some of it was statistical and some of it seemed more forensic, trying to think beyond what-if scenarios. I feel strongly that there was a robust look at the data from both the FDA and also the sponsor. Although folks have referred to this as a small trial, I think for a rare disease setting, cancer setting, there was a large number of events for progression-free survival. Despite the censoring, there was still a large number of events on both arms.

Statistically, even with the varied and many what-if scenarios for changing the results or imputing results for patients, we saw remarkably consistent effect. So I felt that in the context of an open-label trial, I'm able to make the kind of interpretation I want, which is seeing a statistically different value in progression-free survival that I can interpret as probably modest at best, but it seemed reliably interpretable, given the context presented, especially in the context of

the totality of the data, that here we have an unvalidated surrogate and we have overall survival in our hands to continue to interpret what we think this progression-free survival really means for the patient experience, and that was my rationale. Thank you.

DR. MADAN: Thank you, Dr. Shaw.

Dr. Vasan?

DR. VASAN: Neil Vasan. I voted no. Drugging KRAS G12C is certainly a landmark scientific discovery, and there's little question about the activity of sotorasib, but we were asked to comment if the PFS can be reliably interpreted in CodeBreaK 200, and I felt the answer was no. The magnitude of effects is small, statistically significant but not clinically significant, and I do appreciate the rigorous analyses by both the FDA and the applicant.

I do think that this ODAC is an important call for our entire community, our professional organizations, oncologists, industry representatives, patient advocates, and also the

FDA to come up with strategies to take responsibility so that we can mitigate this perception of equipoise, which may have led to biases in this trial. And I think that we as a community have to address this so that we can balance hope with hype for new therapies for our patients. Thank you.

DR. MADAN: Thank you, Dr. Vasan.

Dr. Gradishar?

DR. GRADISHAR: Yes. Bill Gradishar from Northwestern. I voted no, and I share many of the sentiments that Dr. Vasan just expressed. I think this drug is active. It's demonstrated both in this trial and others that it is. It's certainly a more desirable drug, I think, on the whole than receiving docetaxel. That's demonstrated by the toxicity data and the patient's experience.

But I, too, have the same issue with the integrity of the study and the assessments that were made and, actually, the difference in PFS between the arms, as pointed out, I think it may have met what was desired by the trial, but

clinical relevance is a different issue. When the integrity of even that small difference is called into question, despite the 3 hours of statistical gymnastics, I still have as many questions about whether there is anything more than a wash between the two treatment arms with respect to PFS, so I voted no. Thank you.

DR. MADAN: Thank you, Dr. Gradishar.

Mr. Pantelas?

MR. PANTELAS: Yes. Jim Pantelas, patient advocate from Michigan. I voted no, but it's not a question about the drug, or what I feel about the drug, or the importance that I think it offers my community of lung cancer patients and survivors. It's a vote on a very, very narrow topic of imaging review, and what I didn't hear was an explanation for why the image reviews were so vastly different, why the first set of image reviews were so different than the second, or why the second group of reviewers might be better than the first. I hated the question. Thanks.

DR. MADAN: Sorry. I was muted there. I

just said that it's fair to comment on the last point, sir.

Dr. Spratt?

DR. SPRATT: Yes. Thank you to the applicant and the FDA for all the work that was put into this. I think without question, all of us want to help cancer patients improve the way they experience life. We typically obviously quantify that right as improvements in quality or quantity of life or a surrogate of quantity of life. So I guess, first, my assessment of CodeBreaK 200 is as follows.

The drug did not help patients live longer. PFS for second-line therapy, based on the most recent studies I can find, the surrogate threshold effect for survival would have to be less than 0.3. This effect sized was obviously closer to 0.6 or higher, and quality of life, this study was not designed to do a superiority trial for quality of life, and quality of life was not even assessed beyond progression, so we do not know what the global and net long-term quality of life is.

What led to my vote of no -- and I'm surprised not more discussion was on this -- is the fact that PFS-1 was better with the experimental agent and also PFS-2 favored the experimental agent, but there is no difference in overall survival and no further explanation. This leads me to believe with all of the other discussion we've had, that there is likely bias or inaccuracies in the PFS assessment. So there's a high probability of bias with this non-surrogate endpoint. So, unfortunately, I lack the confidence and reliability of the PFS endpoint in CodeBreaK 200. Thank you.

DR. MADAN: Thank you, Dr. Spratt.

Mr. Mitchell?

MR. MITCHELL: Yes. I generally want to associate myself with the remarks of Mr. Pantelas. You didn't ask me if I, as a cancer patient, for example, would like to have this drug available to me. Do I believe, even if they're roughly equal, the fact that it is a drug that's much easier than the control agent for patients? You didn't ask

about what do you think about risk-benefit. You didn't ask whether it should be converted to a full approval. You asked a very narrow question about the conduct of the study, and as Dr. Spratt pointed out, specifically in relationship to the imaging questions and can we put our faith in this study in terms of demonstrating a benefit on progression-free survival. And given the narrow framing of the question, the answer was clearly no, after 3 hours of hard, thoughtful, long discussion, so I voted no.

DR. MADAN: Thank you, Mr. Mitchell.

Dr. Hoffman?

DR. HOFFMAN: Yes. I voted no. I was very strict in my interpretation of the question informing that vote, as we were asked to be, and I do applaud the sponsor for making great efforts to look at the worst-case scenarios to address the concerns that had been raised in the statistical analysis.

I guess my thought as a clinician is that even in the worst-case scenario, if there's

absolutely no difference between sotorasib and docetaxel in terms of efficacy, I still would hope that sotorasib could remain as an option for patients in that clinical setting because probably many of them, or if not most, would choose an oral-targeted drug, or the speed of activity, and so on. So much of what Dr. Johnson said, I totally agree with in terms of being an active clinician, so I would hate to see the drug not continue to be available. But from the standpoint of the strict question, I felt that I needed to vote no on the basis of what we've heard for the last several hours. Thank you.

DR. MADAN: Thank you, Dr. Hoffman.

Dr. Gulley?

DR. GULLEY: Yes. I applaud the FDA for bringing up this important question about the reliable interpretation of CodeBreaK 200. This really was a complicated issue and best discussed in an open forum after evaluation of all the data. I applaud the sponsor for careful, clear analysis and also applaud the tone of the meeting to bring

up these issues in a transparent and unbiased manner.

Clearly, this is an active and, I might add, FDA-approved agent, but I struggled with this narrow question, as I think most of the potential biases can be reliably assessed. I did vote no, but I would say that the early dropout in the docetaxel arm, I think the potential biases could be assessed here. The sensitivity analysis looked good, and the baseline characteristics were not favoring approval of the experimental arm, so I think that was ok. But where I really had issues were with the 19 patients in the docetaxel arm that crossed before progressive disease was evaluated in the blinded radiology review.

So I was glad that there was a hundred percent re-read for the scans for this analysis, and I wasn't worried about the interval censoring analysis because the hazard interval was the same, and the median PFS is really a very arbitrary single point in that Kaplan-Meier curve that should be de-emphasized in relation to the hazard ratio,

which covers the entire curve.

I would also just say that I wouldn't characterize a 34 percent decrease in the risk of a progressive disease or death as marginal; however, when there are biases, then one has to look at the the whole picture. Also, the start of the new anti-cancer therapy seems to be ok from a sensitivity analysis, but the early crossover with no progressive disease and the BICR analysis in those 19 patients, who also appeared to have a better prognosis based on the FDA analysis, that was where I felt like I couldn't overcome the potential issues with that bias. I couldn't address that effectively enough. Thank you.

DR. MADAN: Okay. Thank you, Dr. Gulley.

With that, I'll just briefly summarize. I think despite votes on both sides, yes and no, I think there was relative unanimity in terms of the lamenting a little bit of the narrow focus of the question, which was really focused on this CodeBreaK study and the specific reliability of the data.

For those people who interpreted the data as being relatively unreliable, they voted 10 versus 2. The opinions, though, were pretty much along the lines of the questions that have been raised about the blinded central review of the radiology readouts, the early crossovers and how that contributed to early dropouts, and general study integrity. Questions were also raised a little bit about the inconsistent findings between progression-free survival as it read out with both first- and second-line therapies, and then not translating to an overall survival benefit.

I do think that despite voting no, most of the people expressed optimism that this treatment can be effective, and perhaps we just need more data from a different trial to give a reliable readout on that. For the two people who voted yes and thought this was reliable, it was primarily based on the desire to really interpret the study as it was intended, with progression-free survival as the primary readout, and the thinking that perhaps all these statistical permutations we went

through today kind of convoluted the initial positive finding. So I think those were the comments that predominated in this discussion, and really through the course of the day.

I do want to thank members of the FDA and the sponsor for respecting the committee and presenting a very statistically complicated, nuanced discussion in ways that us and the public can understand, as well as the respectful discourse throughout. A lot of time went into this on the FDA side and the Amgen side, and I think you guys did a great job of presenting the data so it can be understood and plainly available.

I also want to thank the people who spoke in the open public forum. There were a lot of patients who spoke as well, and I thought that they spoke their cases eloquently and shared very personal stories at times, which I think were helpful for the committee to hear.

So with that, I would like to just make sure that there are no additional comments from the FDA before we formally adjourn.

DR. SINGH: I think, Dr. Madan, both Dr. Pazdur and myself, if I may go, and then I'll allow Dr. Pazdur to close.

DR. MADAN: Go ahead.

DR. SINGH: We deeply appreciate the committee, not only the vote but the discussion, and we do hear the conflict in your thought process around the vote about totality and the desire to keep sotorasib on market as an option for patients. We stated in our FDA presentation twice that it is not our intent to immediately withdraw a drug that has a, quote/unquote, "failed confirmatory trial." It is under accelerated approval, and there are multiple pathways available to us, and we are not making this move to withdraw the drug from the market based on these results. We have not indicated that, and we are taking, again, into account your discussion.

Today's discussion was recorded, and it sounds like we have a call to action, in fact, to discuss moving forward conduct and mitigation strategies in open-label trials. I appreciate all

the comments. Both the FDA and the applicant put immense effort into this, and we really appreciate the committee's thoughtful discussion today. And with that, I will defer to Dr. Pazdur.

Thank you, Dr. Madan.

DR. PAZDUR: Here again, I'd like to thank everybody. This was a great discussion, and also the patients that participated in the open public hearing.

I do want to follow up with a comment that I made earlier and that was echoed by Dr. Neil Vasan. We have particular interest in the integrity of the clinical trial system, and it is quite bothersome to me and the agency, in general, when we see unidirectional dropout on clinical trials to this degree. This is something that we have to address in the oncology community, particularly. Why? Because we do have, generally, unblinded trials.

So we will be following up with this with various professional groups and various external symposiums to have further discussion on this entire issue. Here again, I think it's very

important that investigators really enter a clinical trial and have a commitment to enrolling patients, and not use a trial to get access to drugs, and then say, "Well, at the end of the randomization process, if somebody didn't get the drug, I might not proceed with the trial."

We have seen this in other trials in oncology and, fortunately, in those trials, the problem was obviated by a big effect on overall survival; but here again, that does not mitigate the problem in general. So we as an oncology community have to address this issue. No amount of statistical machinations will address a poorly conducted trial, so we really have to address this from a long-term perspective.

If people are agreeing to go on a study, if investigators are willing to participate in a trial, they have to commit to really proceeding with the way the trial was written, and I think that this is an important conversation that we have to have in the oncology community because, here again, we have been seeing this, and this is a

great deal of concern that I have, and the agency, as we move forward in the evaluation of oncology agents. And I will leave that at that, and I thank everybody, but this will be a continuing discussion that we will have.

Adjournment

DR. MADAN: Okay. I think with that, we will now adjourn the meeting. Thank you, everyone, for taking part.

(Whereupon, at 3:03 p.m., the meeting was adjourned.)