FDA Briefing Document

NDA# 215010

Drug name: Gefapixant

Applicant: Merck Sharp & Dohme Corp.

Pulmonary-Allergy Drugs Advisory Committee Meeting

November 17, 2023

Division of Pulmonary, Allergy, and Critical Care – Office of Immunology and Inflammation

DISCLAIMER STATEMENT

The attached package contains background information prepared by the Food and Drug Administration (FDA) for the panel members of the Advisory Committee. The FDA background package often contains assessments and/or conclusions and recommendations written by individual FDA reviewers. Such conclusions and recommendations do not necessarily represent the final position of the individual reviewers, nor do they necessarily represent the final position of the Review Division or Office. We have brought the New Drug Application (NDA) for gefapixant, an oral tablet, for the treatment of refractory or unexplained chronic cough in adults to this Advisory Committee in order to gain the Committee's insights and opinions, and the background package may not include all issues relevant to the final regulatory recommendation and instead is intended to focus on issues identified by the FDA for discussion by the Advisory Committee. The FDA will not issue a final determination on the issues at hand until input from the Advisory Committee process has been considered and all reviews have been finalized. The final determination may be affected by issues not discussed at the Advisory Committee meeting.

Table of Contents

Table of Tables

Table of Figures

Glossary

AC	Advisory Committee
ACCP	The American College of Chest Physicians
AE	adverse event
ANCOVA	analysis of covariance
BD	Briefing Document
BID	twice daily
BRF	Benefit-Risk Framework
CC	chronic cough
CDER	Center for Drug Evaluation and Research
CI	confidence interval
COA	clinical outcome assessment
CSD	cough severity diary
CSR	clinical study report
FAS	full analysis set
FDA	Food and Drug Administration
GERD	gastroesophageal reflux disease
LCQ	Leicester Cough Questionnaire
MI	multiple imputation
MMRM	mixed model repeated measures
NDA	new drug application
OR	odds ratio
PGIC	Patient Global Impression of Change
PRO	patient-reported outcome
RCC	refractory chronic cough
SAP	Statistical Analysis Plan
SUI	stress urinary incontinence
TEAE	treatment-emergent adverse event
UACS	upper airway cough syndrome
UCC	unexplained chronic cough
VAS	Cough Severity Visual Analog Scale

1 Executive Summary/Draft Points for Consideration by the Advisory Committee

1.1 Purpose/Objective of the AC Meeting

The Food and Drug Administration (FDA) is convening the Pulmonary-Allergy Drugs Advisory Committee (PADAC) meeting to discuss whether the clinical trial data for gefapixant, an antagonist of the P2X3 receptor, demonstrate a clinically meaningful treatment benefit to support the proposed indication of treatment of refractory or unexplained chronic cough in adults.

1.2 Context for Issues to Be Discussed at the AC

Chronic cough is defined as a cough that is present for more than 8 weeks. While it is not a life-threatening condition, CC can have a negative impact on quality of life. CC is a common reason for patients to seek clinical care and may occur in 5% to 10% of adult patients (Gibson et al. 2016). Patients who have been diagnosed with conditions that could cause CC (e.g., upper airway cough syndrome (UACS), gastroesophageal reflux disease (GERD), asthma, chronic obstructive pulmonary disease (COPD)) and whose cough does not resolve with appropriate treatment of the underlying condition, are considered to have refractory chronic cough (RCC). Patients with no underlying etiology for cough are considered to have unexplained chronic cough (UCC). For simplicity, the briefing document will refer to RCC and UCC collectively as chronic cough (CC).

The underlying cause of CC is unclear and continues to be investigated. Increased airway sensitivity to noxious stimuli, which activate sensory C-fibers of the vagus nerve leading to initiation of a cough reflex, may play a role in the pathogenesis of CC. Gefapixant is hypothesized to act by inhibiting purinergic receptor P2X3, an ATP-gated ion channel found on sensory C-fibers. Through antagonism of P2X3, gefapixant may ameliorate increased sensitivity to noxious stimuli, thereby suppressing the cough reflex.

There are currently no FDA-approved therapies for CC, and therefore this is a therapeutic area of unmet need. Along with nonpharmacologic interventions such as speech pathology therapy, a variety of products are used off-label for treatment, including opioids, neuroleptics, and local anesthetics. However, the available treatments carry risks, and the evidence supporting use of off-label treatments is limited. As such, FDA anticipates that a new product approved for CC will be widely used given the prevalence of the condition and lack of therapeutic options.

1.3 Brief Description of Issues for Discussion at the AC

Brief Regulatory History of the Development Program

Gefapixant is a new molecular entity (NME) that is not FDA-approved for any indication. The Applicant, Merck Sharp & Dohme Corp., submitted a new drug application (NDA) for gefapixant tablets (45 mg by mouth twice daily) for the treatment of CC in adults. The NDA was first submitted on December 21, 2020, and was not approved in the first cycle. The primary deficiency was insufficient validation data to support that the VitaloJAK cough counting system—consisting of a wearable VitaloJAK digital recording device that captures cough sounds; the compression algorithm that removes periods of silence and non-cough sounds from the recording; and review by human cough-counting analysts—provided a reliable and accurate assessment of the primary endpoint, cough frequency. In addition, FDA conveyed concerns about the unclear clinical benefit of gefapixant, including the uncertain clinical meaning of the primary endpoint results, and concerns regarding support from

the patient-reported outcomes (PROs). On January 20, 2022, the FDA issued a Complete Response letter outlining the deficiencies to the Applicant.

To address the insufficient validation of the VitaloJAK cough counting system, the Applicant performed a recording compression algorithm validation study and an inter-rater reliability study. The Applicant recounted the cough counts in the two pivotal clinical trials (P027 and P030) using the processing methods evaluated by these validation studies. Results based on the recount cough data were included in the resubmission of the NDA, received by the FDA on June 30, 2023. Upon review, the FDA's Center for Devices and Radiological Health has concluded that the validation package included in the NDA resubmission is sufficient to allow for substantive review and discussion of the clinical cough recount data in the NDA. The validation of the VitaloJAK device and compression algorithm used in the gefapixant program is unique to the gefapixant pivotal trials and in no way implies validation beyond the existing 510(k) device clearance (i.e., this device and algorithm validation is not considered generalizable). The cough count dataset from the original NDA submission was not produced by a validated, reproducible method. For all FDA analysis and results discussions, we present the cough frequency *recount* data as it was produced via a validated process in a blinded manner, ensuring reliability and accuracy of the primary endpoint results.

Key Aspects of the Development Program

CC is a novel therapeutic indication that lacks regulatory precedent, particularly with regard to endpoint selection and interpretation of efficacy results. As described in this briefing document, our assessment of the submitted data is that treatment with gefapixant showed a small reduction in cough frequency, but we have questions whether the effect is clinically meaningful. Because of the small treatment effect and the limited experience with the endpoints used in this program, the FDA requests that the Committee provide their assessment of the clinical effectiveness of gefapixant, specifically the clinical meaningfulness of the observed results related to cough frequency and supportive information from PROs.

The gefapixant program consisted of two 52-week, randomized, double-blind, and placebo-controlled pivotal trials, P030 and P027, in adults with a diagnosis of CC. Both trials compared gefapixant 45 mg twice daily and 15 mg twice daily to placebo twice daily and evaluated cough frequency as the primary endpoint at Weeks 24 and 12 in Trials P030 and P027, respectively. Cough frequency was assessed by the VitaloJAK cough counting system and calculated as the number of cough events over a 24-hour period divided by total duration of the recording (minimum 20 hours) and expressed as cough frequency. The prespecified primary analysis was mixed model repeated measures (MMRM), which the Applicant refers to as longitudinal analysis of covariance. The Applicant is only seeking approval of gefapixant 45 mg; the 15 mg dose did not show a statistically significant reduction in cough frequency compared to placebo. Therefore, our briefing document (BD) will focus on the results for the 45 mg dose group.

Although the NDA resubmission also included clinical data from supplementary trials (P030 China extension, P042, and P043), study design and/or conduct issues limit their utility. P042 enrolled a subpopulation of CC patients who had stress urinary incontinence (SUI) and evaluated cough-induced SUI episodes as the sole efficacy endpoint. P043 enrolled patients with a more recent diagnosis of CC; Leicester Cough Questionnaire (LCQ) total score was the primary efficacy variable. Neither P042 nor P043 assessed cough frequency. Efficacy outcomes in the P030 extension in China were impacted by the COVID-19 pandemic. Based upon our review, the results from these trials do not aid in the interpretation of the meaningfulness of the clinical benefit of gefapixant for CC. Consequently, these trials will not be a focus of our discussion and are included in the Appendix (sections 6.2 and 6.3) for reference.

In the NDA resubmission, the Applicant conducted additional analyses using unvalidated original cough data and validated recount cough data, each analyzed with the prespecified MMRM analysis (described by the Applicant as longitudinal ANCOVA) and a post hoc multiple imputation (MI) and ANCOVA method. We view the recount data as the most reliable dataset because it was generated using a validated algorithm in a blinded manner. Therefore, the prespecified MMRM analysis of the recount data is deemed the primary analysis for efficacy, and our BD focuses on these results. An overview of the results of the primary and key secondary endpoints related to cough frequency in the pivotal trials with the prespecified analysis is shown in Table 1.

Table 1. Overview of Cough Frequency Primary and Secondary (Multiplicity-Controlled) Efficacy Results, Trials P030 and P027 (Full Analysis Set, Recount Data)

Variable	Trial P030 (Week 24)		Trial P027 (Week 12)	
	Placebo	Gefapixant 45 mg	Placebo	Gefapixant 45 mg
N	419	409	222	217
Geometric mean[1] at baseline	20.4	19.4	23.6	18.9
Geometric mean[1] at Week 24 or 12	8.7	7.1	10.6	7.4
Primary endpoint/analysis/p-value				
Relative reduction (%) in geometric mean ratio (95% CI)2		-14.6 (-26.0, -1.5)		-17.0 (-31.5, 0.6)
p-value		0.030		0.057
Median[3] cough frequency at baseline (min, max)	21.3 (0.7, 184)	19.9 (0.2, 230)	26.1 (0.3, 1054)	20.9 (0.2, 399)
Median[3] cough frequency at Week 24 or 12	11.4	7.7	11.6	8.7
Median[3] change from baseline in cough frequency at Week 24 or 12	-8.7	-9.8	-8.9	-10.5
Awake cough frequency[‡]				
N	419	409	222	217
Geometric mean[1] at baseline	26.9	25.2	31.4	25.0
Geometric mean[1] at Week 24 or 12	11.3	9.0	13.8	9.6
Relative reduction (%) in geometric mean ratio (95% CI)[2]		-15.5 (-27.0, -2.3)		-16.3 (-31.2, 1.9)
p-value		0.023		0.076
≥30% reduction from baseline in 24-hr cough frequency[†]				
N[4]	368	347	205	194
n[5] (%)	245 (67)	248 (72)	135 (66)	133 (69)
Odds ratio vs. placebo (95% CI)[6]		1.2 (0.9, 1.7)		1.2 (0.8, 1.8)
p-value		0.188		0.435
Responder/N (%)[7]	245/432 (57)	248/434 (57)	135/232 (58)	133/237 (56)
Estimated difference (%) (95% CI)[8]		0.4 (-6.1, 7.0)		-2.2 (-11.2, 6.8)

Source: adeff.xpt; Table 4-2 in CSR Addenda for P030 and P027, Tables 4-4 and 4-10 in CSR Addenda, Table 14.2-22 in CSR of Trial P030, Tables 4-3, 4-4 in CSR Addenda of Trial P027, validated by the statistical analyst.

[†] Cough frequency was assessed by the VitaloJAK cough counting system and calculated as the number of cough events over a 24-hour period divided by total duration of the recording (minimum 20 hours) and expressed as cough frequency.

[‡] Awake cough frequency assessed by the VitaloJAK cough counting system calculated as the number of cough events over the wake period divided by total duration of the awake recording and expressed as cough frequency. Determination of awake and sleep states was conducted by the VitaloJAK cough analyst using a protocolized method that incorporates the local time of the recording (e.g., 10 pm local time was considered an average sleep time) as well as decreased audio activity in the recording.

[1] Geometric means were used because these frequency data were log-transformed.

[2] Based on MMRM model. The estimated relative reduction (relative to placebo) is calculated by 100 (e^{DIFF} -1). DIFF is the treatment difference in change from baseline at Week 12 or 24 based on the log-transformed data.

[3] Median values were from post hoc analyses on all available observations at specific visits.

[4] N=number of subjects who had baseline values.

[5] n=number of responders at Week 12 or 24.

[6] Based on the logistic regression model with terms of treatment, visit, treatment-by-visit interaction, gender, region, baseline, and baseline-by-visit interaction among subjects who had baseline and post baseline values.

[7] N=number of subjects who had a baseline value. Subjects with nonmissing baseline data but missing Week 24 data are imputed as nonresponders.

[8] Based on the Miettinen and Nurminen method.

Abbreviations: CI, confidence interval; CSR, clinical study report; MMRM, mixed model repeated measures

The primary endpoint was analyzed as the mean change from baseline in the natural log-transformed cough frequency at Week 24 or 12, in Trials P030 and P027 respectively, and characterized as the relative reduction over placebo (the Applicant performed log transformation to address expected skewness of data). It is important to note the large placebo effect in both trials. The prespecified analysis of the recount cough data yielded a statistically significant difference between geometric mean ratios of the proposed 45 mg dose and placebo in Trial P030 only. We note that use of the recount data shifted the p-value to >0.05 for the smaller Trial P027. However, the point estimate for the reduction in cough frequency was similar in both trials, a relative reduction in the geometric mean ratio of -15% to -17% compared to placebo from baseline to Week 24 or Week 12. The observed treatment effect is considerably smaller than the 30% relative reduction that Trials P027 and P030 were powered to detect; we note that P030 had a larger sample size to provide additional power for assessment of the key secondary endpoint of a ≥1.3-point increase from baseline in LCQ total score.

Given the complicated presentation of the data (i.e., log transformation, geometric mean ratio, relative reduction), assessment of the clinical meaning of the results is a challenge. Therefore, we conducted a post hoc analysis of the absolute cough frequency, a more intuitive expression of the primary endpoint, which revealed small differences between treatment groups in the median cough frequency as shown in Table 1. In P030 and P027, the baseline mean and median cough frequencies were roughly 20-30 coughs per hour with an upper range of hundreds of coughs per hour. We question if gefapixant treatment, resulting in a reduction beyond the high placebo response of approximately 1 to 2 coughs per hour, results in a benefit that is perceptible to patients; we ask the AC panel to discuss the clinical meaningfulness of the reduction in cough frequency with gefapixant relative to placebo.

In addition, FDA looked at other endpoints related to cough frequency and performed other analyses to explore whether the reduction in cough frequency would be meaningful to patients.

Prespecified Secondary Endpoints

As shown in Table 1, awake cough frequency mirrored the primary endpoint of 24-hour cough frequency. The responder analysis for ≥30% reduction from baseline in 24-hour cough frequency did not show a statistically significant difference between gefapixant and placebo in either trial.

Subgroup Analyses

We conducted subgroup analyses of baseline demographics and disease characteristics to explore whether particular subgroups of patients were more likely to respond to gefapixant (Figure 10 and Figure 11). We could not identify any subgroup that could be identified prior to treatment in a clinical practice setting.

Patient Global Impression of Change (PGIC)

We also requested post hoc anchor-based analyses using the Patient Global Impression of Change (PGIC)[1] PRO as an anchor scale. The PGIC is a single-item PRO asking patients to describe their cough "now" compared to the start of treatment with seven response options ranging from "very much improved" to "very much worse." Using the PGIC as an anchor, we can explore whether there is a correlation between patient-reported improvement in cough and the change in cough frequency. The Applicant reports numerical differences in "improvement responders" between treatment groups in both P027 and P030. However, both trials showed low correlation between the PGIC and change in cough frequency (Figure 7; Polyserial/Spearman: 0.15/0.32 and

[1] The PGIC is the only PRO measure administered in both studies that would be considered reasonable as an anchor scale.

0.23/0.30 for Trials P027 and P030, respectively). This poor association of cough frequency with PGIC score indicates that the change in cough frequency occurs independently from patient-reported improvement in chronic cough (as captured by PGIC); in other words, patients who reported feeling better per the PGIC were not necessarily those patients who were coughing less.

We ask the Committee to consider these additional endpoints and analyses in their assessment of the clinical meaningfulness of the primary endpoint results.

Patient-Reported Outcomes (PROs)

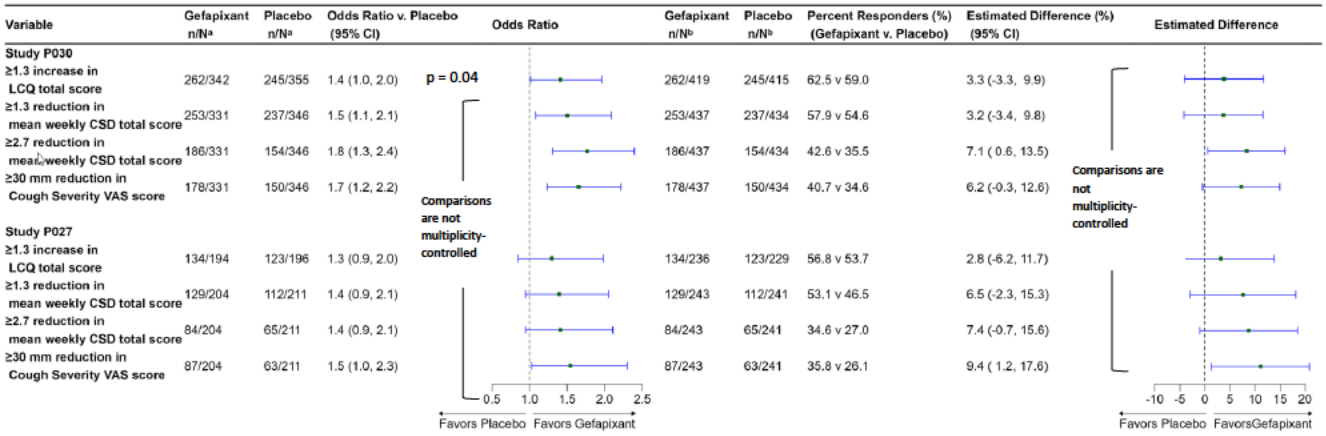
Given the small treatment difference in cough frequency, the direct assessment of patients' experience via PROs may contribute valuable evidence towards understanding the benefit of gefapixant for the symptomatic treatment of CC. FDA values the patient perspective, and as such, encourages the use of fit-for-purpose[2] clinical outcome assessments such as PROs to support regulatory decisions. That said, there should be sufficient qualitative and quantitative validity evidence to support the interpretation that the PRO score(s) reflect the concept(s) of interest[3] within the target context of use. In other words, we expect the PRO to measure what is important to CC patients; the PRO score to be an accurate and reliable measure of the effect that is important to CC patients; and the change in PRO score to be understandable and to correspond to clinically meaningful improvement from the patient's perspective. To determine whether we use a PRO for regulatory decision-making (fit-for-purpose), we consider the strength of the data to support what the PRO measures and the data to support interpretation of the score.

Given the novelty of the indication and lack of regulatory experience with PROs for CC, it was reasonable for the Applicant to collect data from patients using various PROs in the gefapixant program. That being said, we do expect the Applicant to provide data to support that the PROs being relied upon are fit-for-purpose as described above. The Applicant evaluated several PROs in their development program, including the LCQ, Cough Severity Diary (CSD), and Cough Severity Visual Analog Scale (VAS). Odds ratios and differences for the prespecified PRO-based secondary endpoints are shown in Figure 1. We note that similar to the cough frequency, there is a large placebo effect with the PROs. The Applicant proposed odds ratio as their main summary measure for these endpoints. Because odds ratios can be difficult to interpret clinically, we have also displayed differences estimated by the other analysis prespecified by the Applicant.

[2] Fit-for-purpose: A conclusion that the level of validation associated with a tool is sufficient to support its context of use

[3] Concept of interest: The concept is the aspect of an individual's clinical, biological, physical, or functional state, or experience that the PRO assessment is intended to capture or reflect

Figure 1. Results of Prespecified Patient-Reported Outcome-Based Secondary Endpoints (Full Analysis Set)



Source: Statistical analyst and Table 10 and Table 12 in Section 3.1.1.
Gefapixant: gefapixant 45 mg. Change from baseline at Week 24 (Trial P030)/12 (Trial P027).
Odds ratio is based on the logistic regression model with terms of treatment, visit, treatment-by-visit interaction, gender, region, baseline, and baseline-by-visit interaction among subjects who had baseline and post baseline values.
There was a small discrepancy in odds ratio results for ≥2.7 CSD reduction for Trial P027.
Estimated difference is based on the stratified Miettinen and Nurminen method with gender and region.
Abbreviations: CI, confidence interval; CSD, cough severity diary; LCQ, Leicester cough questionnaire; n, number of responders at Week 24 for Trial P030 (Week 12 for Trial P027); N[a], number of subjects with available values at Week 24 for Trial P030 (Week 12 for Trial P027); N[b], number of subjects who had baseline values

Leicester Cough Questionnaire (LCQ)

Based on the testing hierarchy in each trial, the LCQ total score is the only PRO endpoint that achieved statistical significance. As shown in Figure 1, results from P030 showed a statistically significant, but numerically small, increase in the odds ratio for the proportion of patients who had a ≥1.3-point increase from baseline in the LCQ total score (range 3 to 21). While the OR for the LCQ total score in P030 was statistically significant, the proportion of patients who had a ≥1.3-point increase from baseline in the LCQ total score is only 3.3% more than placebo Figure 1.

The LCQ is a 19-question PRO that covers 3 domains related to cough and its impact. The three domains are physical, psychological, and social; the total score (range 3 to 21) is based upon data from all 3 domains with higher scores indicating better health status (see Section 6.4.5). FDA has concerns about whether the LCQ total score is fit-for-purpose. Specifically, concepts evaluated in the social and psychological domains (e.g., embarrassed or worried about cough; cough has interfered with enjoyment of life) can be influenced by factors outside of treatment. In addition, there is a lack of information to determine what change in score is clinically meaningful, and the Applicant did not provide sufficient evidence to support a responder threshold of ≥ 1.3 points (see Section 6.4.2).

To assist with interpretation of the responder analysis on LCQ total score, FDA conducted post hoc analyses of the change from baseline in total score and in each of the individual domains contributing to the LCQ total score. The absolute change from baseline in the LCQ total score compared to placebo was small in relationship to the range of possible scores (0.78 points at Week 24 and 0.35 points at Week 12 in P030 and P027, respectively, out of a possible range of 3 to 21). The physical domain, which is of greater relevance because it more directly assesses cough and its impact, appears to contribute slightly less improvement towards the total score than the other domains (Table 13). Further, FDA conducted a post hoc analysis on change from baseline in the LCQ physical domain score at Week 12 (for P027) and Week 24 (for P030) that showed no difference between gefapixant and placebo. Overall, because of the limitations in interpreting the responder analysis, the small

difference in absolute score, and the lack of effect on the physical domain score, we ask the Committee to consider whether the LCQ results offer insight into the meaningfulness of the primary endpoint results.

Cough Severity Diary (CSD) and Cough Severity Visual Analog Score (VAS)

The treatment effects on proportions of responders based on CSD total score and Cough Severity VAS score were evaluated as secondary endpoints, which were not controlled for multiplicity; these responder analyses demonstrated small numerical increases in the odds ratios as shown in Figure 1. The Applicant did not provide evidence to support that the selected responder thresholds correspond to a clinically meaningful change.

Although the PRO-based endpoints appear to favor gefapixant, the results must be interpreted with caution because 1) the measured absolute differences from placebo in the total score are small and difference in responders between treatment groups is small, 2) the degree of change in PRO scores corresponding to clinically meaningful improvements has not been established, and 3) potential unblinding due to taste disturbance effects (in up to 65% of treated subjects, discussed below) could impact the interpretation of the results, and 4) with the exception of the LCQ responder thresholder analysis in P030, none of these analyses are controlled for multiplicity.

In light of these limitations and uncertainties, we ask the Committee to consider whether gefapixant's small treatment effect across PRO-based endpoints is meaningful to patients. Importantly, we ask the Committee to discuss whether the PRO data provide compelling evidence to inform the key question of whether the small reduction in cough frequency with gefapixant is clinically meaningful to patients.

Safety

The safety profile of gefapixant 45 mg is notable for frequent taste disturbances. In the pivotal trials, disturbance in or loss of taste occurred in up to 65% of subjects in the gefapixant 45 mg dose group and substantially impacted its tolerability, leading to discontinuation of treatment in 14% of subjects (compared to <1% of placebo subjects). The taste disturbance has rapid onset (median 2 days) along with a mean duration of 204 days. The events are generally reversible, resolving in 96% of subjects primarily upon cessation of treatment. Although not a serious risk, this adverse reaction must be weighed against the potential benefit offered by gefapixant for the symptomatic treatment of CC. Beyond posing a safety and tolerability issue, this common side effect has the potential to unblind subjects who experience it, introducing uncertainty to the interpretation of the results, including cough frequency and PRO-based endpoints. Awareness of the frequent occurrence of taste disturbance, which was disclosed in the informed consent and investigator brochure, may have introduced bias for the PRO-based endpoints, based on knowledge of assigned treatment.

Substantial Evidence of Effectiveness

The statutory standard for product approval requires that a drug's effectiveness be established by "substantial evidence." FDA generally requires at least two adequate and well-controlled clinical investigations, each convincing on its own, to establish effectiveness; this is the standard expectation to support a CC indication. It is also well established that the effect must be clinically meaningful. The gefapixant program includes two adequate and well-controlled pivotal trials demonstrating a small reduction in cough frequency and a small increase in responders on the LCQ total score compared to placebo. Exploratory analyses of the other PRO secondary endpoints on CSD and Cough Severity VAS scores show small numerical increases in 'responders' treated with gefapixant compared to placebo.

Statistical significance does not, by itself, indicate whether the detected effect corresponds to a clinically meaningful treatment effect. As such, the question before the Committee is not one of statistical significance, but whether these small treatment effects are clinically meaningful. Conceptually, reducing cough frequency in patients with CC could be considered an important treatment benefit if the results are robust. However, in the gefapixant program, the clinical relevance of the reduction in cough frequency is challenging to assess due to the large variability in baseline cough frequency; the high placebo response; the small magnitude of the treatment response relative to placebo; and the uncertainties surrounding the interpretation of the limited supporting evidence from PROs, which also showed a small treatment effect.

Gefapixant is intended to treat a common, chronic, symptomatic condition and is neither curative nor disease-modifying and does not affect major morbidity/mortality. In this context, it is important to determine that the therapy offers a meaningful impact on how patients feel relative to how they feel when administered placebo. Given the novel indication with novel endpoints, lack of approved therapies, and stakeholder interest in this area of development, we request the Committee's input on whether the submitted data establish evidence of a clinically meaningful benefit of gefapixant for the proposed indication.

1.4 Draft Points for Consideration

- Discuss the evidence of effectiveness for gefapixant for the treatment of refractory or unexplained chronic cough in adults. Specifically address the following:
  - The small reduction in cough frequency compared to placebo and the clinical meaningfulness of the reduction in cough frequency
  - The observed results from patient-reported outcomes (PROs) and whether these results provide compelling evidence to inform the clinical meaningfulness of the reduction in cough frequency
  - Potential unblinding of patients due to taste disturbance and its impact on interpretation of cough frequency and PRO results
- Discuss the overall benefit/risk assessment of gefapixant for the treatment of refractory or unexplained chronic cough in adults, a symptomatic condition.
- Discuss whether the evidence demonstrates that gefapixant provides a clinically meaningful benefit to adult patients with refractory or unexplained chronic cough, given the small reduction in cough frequency and results from PROs.
  - If you conclude that there is insufficient evidence of a clinically meaningful benefit, describe the evidence that could be collected to show a benefit that is clinically meaningful.

2 Introduction and Background

2.1 Background of the Condition/Standard of Clinical Care

Chronic cough is defined as a cough that is present for more than 8 weeks. While it is not a life-threatening condition, it can have a negative impact on quality of life. Studies have shown that chronic cough is a common reason for patients to seek clinical care, with a global prevalence of 5% to 10% of adult patients. (Gibson et al. 2016) In a survey of over 10,000 patients presenting to cough specialist clinics, roughly two-thirds were female with a mean age of 55 years (Morice et al. 2014a). Many patients undergo extensive clinical investigations, as well as empiric therapy trials, to diagnose and treat chronic cough. However, in many patients, chronic cough persists despite these interventions. The natural history of CC is poorly understood; in some patients, it presents as daily cough lasting years or decades, and in others, the course is relapsing and remitting.

Patients who have been diagnosed with conditions that could cause chronic cough (e.g., UACS, GERD, asthma, COPD) but whose cough does not resolve with appropriate treatment of the underlying condition are considered to have RCC. Patients with chronic cough who do not have any underlying conditions suspected of causing the cough are considered to have UCC.

The underlying cause(s) of CC are unclear and continue to be investigated. Cough is a vital protective reflex that serves to prevent aspiration and facilitate airway clearance, and the cough reflex is mediated by a complex and redundant neurophysiologic process. A variety of nociceptors (including purinergic receptors such as P2X3) and mechanoreceptors have been implicated as "cough receptors" in the respiratory mucosa, which respond to both intrinsic and extrinsic noxious stimuli, as well as mechanical stimulation. Once activated, the receptors signal via vagal afferent nerve endings located in the larynx, trachea, carina, and large intrapulmonary bronchi. Sensory information is processed by the brainstem and also involves higher brain (cortical and subcortical) controls (Canning et al. 2014). As such, higher brain functions can serve to both inhibit and voluntarily activate the cough reflex. There is a body of scientific literature supporting the contribution of heightened sensitivity of the afferent limb of the cough reflex (i.e., airway sensitivity to noxious stimuli) to the pathogenesis of CC (Morice et al. 2014b). As an antagonist of the P2X3 receptor, which is expressed on sensory neurons in the afferent limb of the cough reflex, gefapixant is hypothesized to ameliorate this increased sensitivity to noxious stimuli, which could suppress the cough reflex .

There are currently no FDA-approved therapies for CC. A variety of products are used off-label for treatment, including opioids, neuroleptics, and local anesthetics. The American College of Chest Physicians (ACCP) recommends empiric treatment with gabapentin for patients with UCC (Gibson et al. 2016). The European Respiratory Society (ERS) recommends either low-dose morphine, gabapentin, or pregabalin for treatment of chronic cough (Morice et al. 2020). Additional off-label therapies include but are not limited to benzonatate, codeine, dextromethorphan, and guaifenesin. If approved, gefapixant would likely be widely used as a chronic therapy in patients for whom chronic cough is an established diagnosis.

2.2 Product Information

Gefapixant is a new molecular entity and first-in-class P2X3 inhibitor. Gefapixant is formulated as a 45 mg tablet, and the proposed dose is 45 mg taken by mouth twice daily. Gefapixant is not marketed or approved in the United States for any indication.

The proposed mechanism of action of gefapixant is antagonism of purinergic receptor P2X3, which is an ATP-gated ion channel found on peripheral sensory nerves of dorsal root ganglia and expressed in upper and lower airway fibers. When stimulated by ATP released in response to inflammation or irritants in airway tissue, P2X3 mediates reflex responses including cough. P2X3 inhibitors, such as gefapixant, are hypothesized to suppress the cough reflex .

2.3 Regulatory History

Gefapixant was studied under Investigational New Drug (IND) 123007, opened on September 5, 2014. Key interactions between the Applicant and the FDA during clinical development are summarized in Table 2.

Table 2. Summary of Key Regulatory History

Date	Interaction	Highlights
6/19/2017	End of Phase 2 Meeting	• Alignment on 24-hour cough frequency as an acceptable primary endpoint. • Alignment on pivotal trials' duration with one trial evaluating efficacy at 3 months and one trial at 6 months. The Applicant initially proposed evaluating efficacy at 3 months in both trials, and the FDA recommended a longer main study period; the decision to lengthen only one trial was at the Applicant's discretion. • FDA conveyed concerns about the CSD and Cough Severity VAS (consistent with the concerns detailed in Section 3.1.3.5.4), and the proposed CSD responder analysis. Alignment on how the gefapixant development program would address these issues was not reached.
9/28/2017	Type C Meeting	• FDA conveyed concerns about the use of the CSD total score. • FDA recommended that the Applicant use multiple anchor scales to inform the threshold(s) for meaningful change in PRO scores, including a patient global impression of severity (PGIS) scale, given that the patient global impression of change (PGIC) scale requires recall over a long period of time and is thus subject to recall error.
7/1/2020	Pre-NDA Meeting	• FDA noted topline study results showed a modest treatment effect in P030 and P027 and lacked supportive evidence from later timepoints to demonstrate durability of response. • FDA expressed concerns with the LCQ total score, including content validity and the selected responder threshold. • FDA noted the high rate of treatment discontinuation due to adverse events and potential impact on the benefit-risk assessment.
12/21/2020	Initial NDA submission	• The user fee goal date was extended based on a major amendment received June 21, 2021, regarding additional validation data to support the accuracy and reliability of the VitaloJAK system in assessing the primary endpoint.
1/20/2022	Complete Response Letter issued	• NDA deficiencies included insufficient validation data to support that the VitaloJAK cough counting system provides a reliable and accurate assessment of cough frequency, the primary endpoint. • FDA expressed concerns that the primary endpoint results were numerically modest and of unclear clinical significance and stated that secondary endpoint support would be an important factor in evaluating efficacy. • FDA noted the general lack of support from secondary endpoints and questioned whether the PRO tools utilized in the pivotal trials are fit-for-purpose. • FDA questioned the content validity of the LCQ total score and the proposed threshold for meaningful within-patient change in the LCQ total score.
3/7/2022; 7/12/2022; 1/19/2023	Type A, C, and pre-NDA meetings	• Agreement on the design of the validation studies needed to support the VitaloJAK cough counting system and algorithm

Source: Clinical reviewer.

Abbreviations: CSD, Cough Severity Diary; FDA, Food and Drug Administration; NDA, new drug application; LCQ, Leicester Cough Questionnaire; PRO, patient-reported outcome

3 Summary of Issues for the AC

3.1 Efficacy Issues

The Applicant's clinical program for gefapixant included two pivotal trials P030 and P027 in patients with CC. The primary efficacy variable was cough frequency as measured by the VitaloJak system. Our assessment of the efficacy data is that treatment with gefapixant showed a small reduction in cough frequency, but we have questions about whether the effect is clinically meaningful. The Applicant also included PROs in the pivotal trials as supportive information. Results from the PROs showed small increases in the proportion of responders in patients treated with gefapixant compared to placebo, but as described in Sections 3.1.3.5.2 and 3.1.3.5.4, there are issues that need to be considered with respect to the PROs. We seek the Committee's assessment of the clinical meaningfulness of the observed reduction in cough frequency with gefapixant, given the small reduction in cough frequency and supportive information from patient-reported outcomes. In this efficacy section, we provide more details of the efficacy findings to inform your discussion.

3.1.1 Sources of Data for Efficacy

The Applicant submitted results from two pivotal trials, P030 and P027, as well as two additional phase 3b clinical trials, P042 and P043, and an extension of P030 in China. Details of these trials are provided in Table 3.

Table 3. Efficacy and Safety Trials Submitted to Support Registration

Trial Identity	Trial Population	Trial Design	Number Treated, Regimen	Primary Endpoint	No. of Sites and Countries
Phase 3 Pivotal Efficacy and Safety Trials					
P030 P030MK7264 NCT03449147	Adults with RCC or UCC	52-week, R, DB, PC, PG	Total treated: 1314 • Gef 45 mg: 439 • Gef 15 mg: 440 • Pbo: 435	24-hour cough frequency at Week 24	175 sites in 20 countries
P027 P027MK7264 NCT03449134	Adults with RCC or UCC	52-week, R, DB, PC, PG	Total treated: 730 • Gef 45 mg: 243 • Gef 15 mg: 244 • Pbo: 243	24-hour cough frequency at Week 12	156 sites in 17 countries
Phase 3b Supplementary Efficacy Trials					
P042 P042MK7264 NCT04193176	Adult females with stress urinary incontinence and RCC or UCC	12-week, R, DB, PC, PG	Total treated: 375 • Gef 45 mg: 185 • Pbo: 190	Daily episodes of cough-induced stress urinary incontinence at Week 12	90 sites in 12 countries
P043 P043MK7264 NCT04193202	Adults with recent-onset (<12 months) RCC or UCC	12-week, R, DB, PC, PG	Total treated: 415 • Gef 45 mg: 206 • Pbo: 209	LCQ total score at Week 12	91 sites in 12 countries
Phase 3 Supplementary Safety Trial					
P030 P030MK7264 China specific extension	Adults with RCC or UCC	52-week R, DB, PC, PG	Total treated: 160 • Gef 45 mg: 66 • Gef 15 mg: 27 • Pbo: 67	24-hour cough frequency at Week 24	20 sites in China

Source: Clinical reviewer. All treatment doses were given twice daily.

Abbreviations: DB, double-blind; Gef, gefapixant; LCQ, Leicester Cough Questionnaire; Pbo, placebo; PC, placebo-controlled; PG, parallel group; R, randomized; RCC, refractory chronic cough; UCC, unexplained chronic cough

Trials P030 and P027 represent the primary sources of efficacy and safety data submitted by the Applicant to demonstrate substantial evidence of safety and effectiveness for the proposed indication and thus are the focus of FDA's review and BD. While the FDA also evaluated the new clinical trial data from P030 extension, P042, and P043 that were included in the NDA resubmission, these trials are limited in their ability to inform efficacy conclusions due to study design or conduct issues.

P042 was a 12-week placebo-controlled trial to assess the efficacy and safety of gefapixant in reducing the frequency of cough-induced stress urinary incontinence (SUI) in female subjects with RCC or UCC. The enrollment criteria were intended to define a population with "cough-induced" SUI, but this is not recognized as a distinct subpopulation given that SUI has multiple triggers, such as sneezing, laughing, and other activities causing increased abdominal pressure. Likewise, FDA does not consider "cough-induced" SUI to be a standalone indication. Clinical trials for SUI are expected to evaluate the change in all-cause incontinence episodes as a coprimary endpoint alongside a fit-for-purpose PRO. In this trial, however, no secondary efficacy endpoints were assessed to facilitate the clinical interpretation of the primary endpoint. Cough frequency was not assessed in this trial. Due to the inherent trial design limitations, the results of P042 do not aid in the clinical interpretation of efficacy for gefapixant.

P043 was a 12-week placebo-controlled trial to assess the efficacy and safety of gefapixant for recent-onset CC; the major design differences in P043 from the pivotal trials are the restricted population (recent onset CC, within 1 year), shorter trial duration (12 weeks of treatment), and primary endpoint (LCQ total score). Cough frequency was not assessed in the trial. Based upon our review, the results from P043 do not aid in the interpretation of the meaningfulness of the clinical benefit of gefapixant for CC.

For completeness, the protocols and results of Trials P042 and P043 are described briefly in Section 6.3.1. The China-specific P030 trial was an extended enrollment period of protocol P030 to fulfill local regulatory requirements. Due to impacts of the COVID-19 pandemic on efficacy outcomes, this trial was not included in the efficacy analysis and contributes only to the safety assessment.

3.1.2 Protocol Review of Pivotal Trials P030 and P027

The pivotal trials, P030 and P027, had many similar design features. The following unified protocol review will summarize the common trial design features and highlight notable differences between the two protocols.

3.1.2.1 Administrative Information

Titles

P030: A Phase 3, Randomized, Double-Blind, Placebo-Controlled, 12-Month Study to Evaluate the Efficacy and Safety of MK-7264 (gefapixant) in Adult Participants With Chronic Cough

P027: A Phase 3, Randomized, Double-Blind, Placebo-Controlled, 12-Month Study to Evaluate the Efficacy and Safety of MK-7264 (gefapixant) in Adult Participants With Chronic Cough

Dates

P030: March 15, 2018, to August 20, 2020; Clinical Study Report (CSR) completed December 2, 2020

P027: March 14, 2018, to June 5, 2020; CSR completed December 1, 2020

3.1.2.2 Trial Design

Both P030 and P027 were 52-week, multicenter, randomized, double-blind, parallel-group trials to assess the efficacy and safety of gefapixant in subjects with RCC or UCC.

In each trial, after a 2-week screening phase, subjects were randomized 1:1:1 into one of the following treatment groups with all treatments taken orally by tablet twice daily:

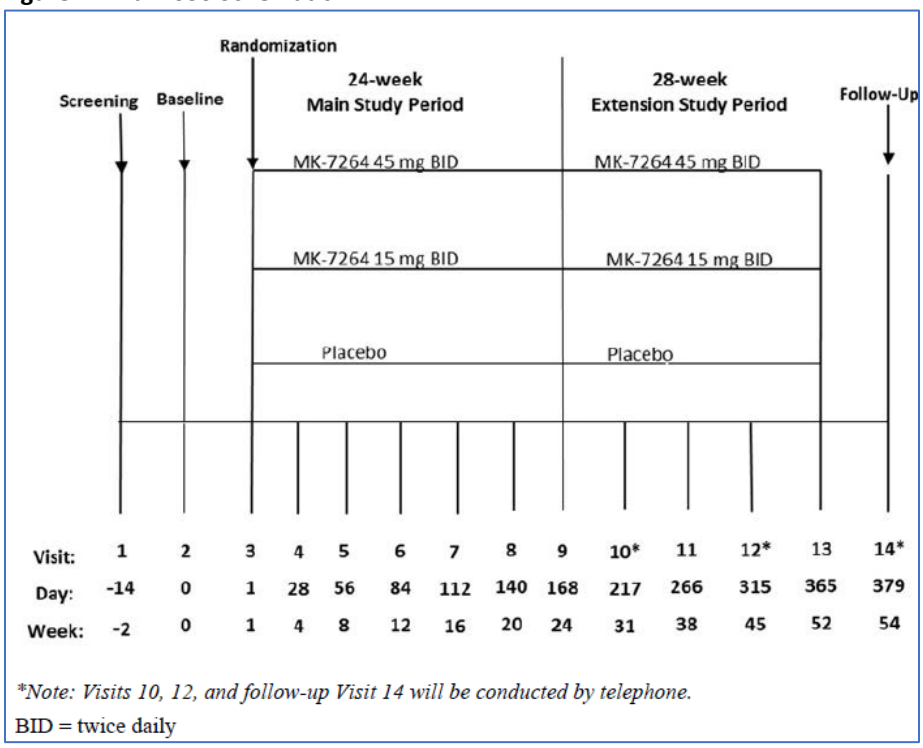
1. Gefapixant 45 mg and placebo matching gefapixant 15 mg
2. Gefapixant 15 mg and placebo matching gefapixant 45 mg
3. Placebo matching gefapixant 45 mg and placebo matching gefapixant 15 mg

P030 randomized 1317 subjects, and P027 randomized 732 subjects. Subjects continued treatment and follow-up through the main study period, at which time efficacy was assessed; this occurred at Week 24 in P030, and at Week 12 in P027. This was followed by an additional 'extension study period' of double-blind treatment until Week 52. PRO efficacy data was collected during the 'extension study period,' but cough frequency data were not.

Screening and informed consent occurred at Visit 1. Baseline data was collected at Visit 2. Randomization occurred at Visit 3 and subjects began study treatment dosing. Following randomization, subjects had study visits approximately every 4 weeks. During the main study period, study visits required subjects to report to the study site on 2 successive days. On the first day, the cough recorder device was attached to the subject, and on the second day, the cough recorder was collected by the study site. Additionally, other safety and efficacy data were collected on the second day of the visit. Cough frequency data was collected approximately every 4 weeks through the end of the main study period (Week 24 in P030 and Week 12 in P027). During the extension period, visits occurred on 1 day, and the cough recorder was not used as cough frequency data were no longer collected.

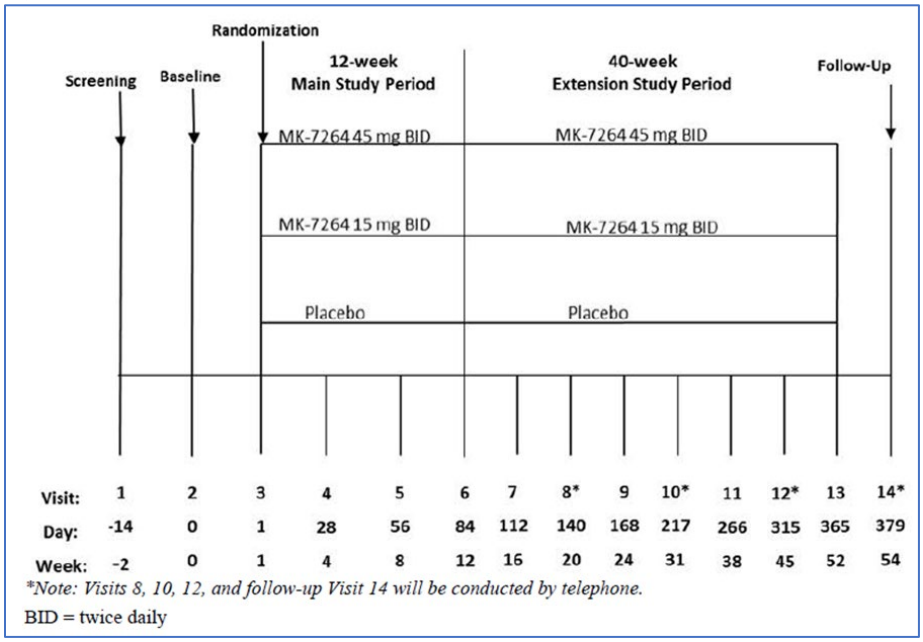
The trial P030 schematic is shown in Figure 2, and the Trial P027 schematic is shown in Figure 3. Overall, the design of P027 was very similar to P030 except for a shorter main study period (12 weeks) and a smaller sample size.

Figure 2. Trial P030 Schematic



Source: Protocol P030MK7264-04; Fig. 1; p. 19.
Abbreviation: MK-7264, gefapixant

Figure 3. Trial P027 Schematic



Source: Protocol P027MK7264-02; Fig. 1; p. 25.
Abbreviation: MK-7264, gefapixant

3.1.2.3 Enrollment Criteria

Key Inclusion Criteria

Male and female participants at least 18 years of age who met all the following:

1. Chronic cough for ≥1 year and a diagnosis of RCC or UCC, defined as:
    a. RCC: Clinical evaluation suggesting a comorbid condition that may be associated with chronic cough, and appropriate diagnostic work-up and at least 2 months of therapy prior to Screening (according to ACCP guidelines) with continued cough despite being on therapy.
    b. UCC: Clinical evaluation of cough per ACCP guidelines without identification of comorbid condition that may be associated with chronic cough.
2. Chest radiograph or computed tomography scan of the thorax (within 5 years of Screening/Visit 1 and after the onset of chronic cough) without detected abnormality contributing to the chronic cough or any other clinically significant lung disease in the opinion of the principal investigator or the sub-investigator.
3. Score of ≥40 mm on the 0 mm to 100 mm Cough Severity Visual Analog Scale (VAS) at both the Screening and Baseline visits.

Key Exclusion Criteria

1. Current smoker or former smokers who had quit within 12 months of Screening or with a smoking history greater than 20 pack-years.
2. FEV1/FVC ratio <60% (spirometry performed within the past year was acceptable if the investigator confirmed that spirometry was done during a period where the participant was clinically stable, e.g., not during an upper respiratory infection).
3. History of upper or lower respiratory tract infection or recent clinically significant change in pulmonary status within 4 weeks of Screening/Visit 1.
4. History of chronic bronchitis, defined as a cough that produces a clinically significant amount of sputum (greater than approximately one tablespoon of phlegm) that occurred every day for at least 3 months in a row, with those periods occurring at least 2 years in a row.
5. Individuals who were currently taking an angiotensin converting enzyme (ACE) inhibitor or had taken an ACE inhibitor within 3 months of Screening.

Medications that were prohibited during the trial are listed in Section 6.1.

The inclusion criteria selected subjects with cough duration of one year or greater, which is a longer duration than the 8 weeks defined by the diagnostic criteria for UCC and RCC. This criterion presumably enriched for a study population with greater diagnostic certainty, as it would be expected that these subjects had previously completed a thorough diagnostic evaluation to rule out other causes and/or treat underlying conditions that could be associated with refractory cough. Additionally, the inclusion criteria selected subjects with a Cough Severity VAS score of 40 mm or greater at baseline; for context, this PRO instrument records patients' assessment of cough severity on a 100-mm scale ranging from "no cough" (0 mm) to "worst cough" (100 mm). There is not an established relationship between the Cough Severity VAS score and cough frequency, the primary endpoint of the gefapixant trials, nor is there convincing evidence that a score of ≥40 mm selects for clinically relevant cough. While there was a criterion for cough severity, there was not an enrollment criterion related to baseline cough frequency, so patients with any baseline cough frequency were eligible for the trial.

3.1.2.4 Efficacy Endpoints

Primary and secondary endpoints were analyzed at Week 24 in Trial P030 and at Week 12 in Trial P027.

Primary Endpoint

- Twenty-four-hour cough frequency at Week 24/12:
  - Cough frequency data were collected over a 24-hour period and expressed as coughs/hour
  - Differences from placebo were expressed as estimated relative reduction
  - The primary endpoint was evaluated for each dose level (gefapixant 45 mg BID and 15 mg BID)

The gefapixant program is one of the first clinical development programs for treatment of CC, so there is limited experience with endpoint selection to evaluate and establish the efficacy of treatments for this condition. It should be noted that selection of endpoints to demonstrate efficacy is intrinsically linked to the proposed drug's mechanism of action, anticipated impacts on clinical outcomes, and the specific target patient population. Typically, efficacy endpoints for a trial evaluating a treatment for a symptomatic condition should measure an improvement in symptom(s) that occur most often in and are most impactful to patients with the condition. In the absence of regulatory experience with PROs for a novel CC indication, cough frequency was a reasonable primary endpoint in the setting of phase 2 trial results showing a cough frequency reduction of roughly 30% relative to placebo; the FDA and Applicant did not prospectively identify the within-patient change in 24-hour cough frequency considered clinically meaningful by patients with CC.

Secondary Endpoints

The prespecified secondary endpoints in the fixed sequence testing hierarchy for Trials P030 and P027 are listed in Table 4. Given the novelty of the indication and lack of regulatory experience with PROs for CC, it was reasonable for the Applicant to collect data from patients using various PROs in the gefapixant program. That being said, we do expect the Applicant to provide data to support that the PROs being relied upon for regulatory consideration are fit-for-purpose. During development, FDA provided feedback to the Applicant on the proposed PROs. An overview of FDA's guidance regarding development of PROs and PRO-based endpoints is provided in Section 6.4.1.

Table 4. Multiplicity Hierarchy for Primary and Secondary Endpoints, P030 and P027

	P030	P027
1	Gefapixant 45 mg is superior to placebo in reducing 24-hour cough frequency at Week 24	Gefapixant 45 mg is superior to placebo in reducing 24-hour cough frequency at Week 12
2	Gefapixant 45 mg is superior to placebo in reducing awake cough frequency at Week 24	Gefapixant 15 mg is superior to placebo in reducing 24-hour cough frequency at Week 12
3	Gefapixant 45 mg is superior to placebo on the proportion of participants with a ≥1.3-point increase from baseline in LCQ total score at Week 24	Gefapixant 45 mg is superior to placebo in reducing awake cough frequency at Week 12
4	Gefapixant 45 mg is superior to placebo with respect to the proportion of participants with a ≥30% reduction from baseline in 24-hour cough frequency at Week 24	Gefapixant 45 mg is superior to placebo with respect to the proportion of participants with a ≥30% reduction from baseline in 24-hour cough frequency at Week 12
5	Gefapixant 15 mg is superior to placebo in reducing 24-hour cough frequency at Week 24.	Gefapixant 15 mg is superior to placebo in reducing awake cough frequency at Week 12

	P030	P027
6	Gefapixant 15 mg is superior to placebo in reducing awake cough frequency at Week 24	Gefapixant 15 mg is superior to placebo with respect to the proportion of participants with a ≥30% reduction from baseline in 24-hour cough frequency at Week 12
7	Gefapixant 15 mg is superior to placebo on the proportion of participants with a ≥1.3-point increase from baseline in LCQ total score at Week 24	
8	Gefapixant 15 mg is superior to placebo with respect to the proportion of participants with a ≥30% reduction from baseline in 24-hour cough frequency at Week 24	

Source: Section 10.8 for Multiplicity in trial protocols for P030 and P027.
Abbreviation: LCQ, Leicester Cough Questionnaire

Other secondary endpoints included:

- Proportion of participants with a ≥1.3-point reduction from baseline in mean weekly CSD total score at Week 24 / 12
- Proportion of participants with a ≥2.7-point reduction from baseline in mean weekly CSD total score at Week 24 / 12
- Proportion of participants with a ≥30 mm reduction from baseline in Cough Severity VAS score at Week 24 / 12

3.1.2.5 Efficacy Assessments

Cough Frequency Assessments

Cough frequency was determined using the VitaloJAK system. In this system, cough sounds were recorded by the VitaloJAK cough recorder over a 24-hour period (minimum 20 hours). The recording was uploaded to the database, then silent periods and non-cough sounds were removed from the recording using a proprietary compression algorithm. A cough analyst then manually counted coughs from the compressed recording, with the output being the cough count over the duration of the recorded period. The recorded period was intended to be 24 hours in duration, but in use some recordings were longer or shorter (recordings of duration <20 hours were treated as missing data). The VitaloJAK cough analysts determined awake and sleep states using a protocolized method that incorporates the local time of the recording (e.g., 10 pm local time was considered an average sleep time) as well as decreased audio activity in the recording. Ideally, cough frequency would have been captured through the end of the 52-week treatment periods; however, the FDA recognized the burdensome nature of this assessment to subjects in the trials. Limiting cough frequency assessments to the main study periods was ultimately at the Applicant's discretion.

To address the insufficient validation of the VitaloJAK cough counting system outlined in the Complete Response Letter, the Applicant performed a recording compression algorithm validation study and an inter-rater reliability study. The Applicant recounted the cough counts in the two pivotal clinical trials (P027 and P030) using the processing methods evaluated by these validation studies. Results based on the recount cough data were included in the NDA resubmission. Upon review, the FDA's Center for Devices and Radiological Health has concluded that the validation package included in the NDA resubmission is sufficient to allow for substantive review and discussion of the clinical cough recount data in the NDA. The validation of the VitaloJAK device and

compression algorithm used in the gefapixant program is unique to the gefapixant pivotal trials and in no way implies validation beyond the existing 510(k) device clearance (i.e., this device and algorithm validation is not considered generalizable). The cough count dataset from the original NDA submission was not produced by a validated, reproducible method. For all FDA analysis and results discussions, we present the cough frequency *recount* data as it was produced via a validated process in a blinded manner, ensuring reliability and accuracy of the primary endpoint results.

3.1.2.5.1 PRO Assessments

PROs were collected electronically. Subjects were trained in the use of the electronic PRO system at screening, and compliance with PRO completion was monitored by the investigator. PROs were collected over the 52-week treatment period (i.e., both the main study period and the extension study period), but prespecified efficacy endpoint analyses were conducted only during the main study period. The frequency of PRO collection varied with each specific instrument and was uniform in both trials.

- LCQ (see Copy of the LCQ) was collected at in-person site visits (Weeks 0, 4, 8, 12, 16, 20, 24, 38, and 52).
- CSD (see Copy of the CSD) and Cough Severity VAS (see Copy of the Cough Severity VAS) were collected daily from Weeks 1 to 24, and at the site visits during Weeks 38 and 52.
- PGIC (see Copy of the PGIC), a single item to assess the change in patient's cough compared to the start of the treatment with seven response options ranging from "very much improved" to "very much worse," was collected at Weeks 12 and 24 only.

For ease of reference, each PRO instrument is described in detail in Section 3.1.3 alongside the presentation of its relevant results.

3.1.2.6 Safety Assessments

Safety and tolerability endpoints were assessed by clinical evaluation of AEs and other study parameters, including vital signs, physical examination, and standard laboratory safety tests.

3.1.2.7 Statistical Analysis

The sample size calculations were based on the primary and key secondary efficacy endpoints, with assumptions based on phase 2 trial P012 data. In both trials P030 and P027, the primary endpoint was powered at >99% for the comparison between gefapixant 45 mg and placebo assuming 30% relative reduction to placebo. In trial P030 only, the ≥1.3 LCQ endpoint was powered at 81% for the comparison between gefapixant 45 mg and placebo assuming proportion of LCQ responders at Week 24 was 54% and 75% in placebo and gefapixant 45 mg, respectively. These calculations yielded a total sample size of 720 for P027 and 1290 for P030, respectively.

To strongly control Type-I error rate for the primary and key secondary endpoints in the multiple testing hierarchy, a step-down testing procedure was applied in the order specified above (Table 4) for both trials. Each hypothesis was formally tested only if the preceding one was significant at α=0.0499 level (an α-spending of 0.0001 was applied to the two-sided Type I error rate of 0.05 for the primary and secondary hypotheses due to the interim analysis when approximately 40% of participants either completed, or discontinued before completing, the main study period).

The Applicant specified the full analysis set (FAS) consisting of all randomized participants who took at least one dose of study treatment as the primary population for the analysis of efficacy in both trials. For endpoints that were measures of change from baseline, the Applicant included subjects who had baseline and at least one post-baseline measurement for inclusion in the analysis of each specific endpoint.

The Applicant considered that the 24-hour cough frequency may be distributed wide and skewed to the right; therefore, the cough frequency data were log-transformed to achieve a normal distribution, an underlying assumption for the primary statistical model for cough frequency. The prespecified analysis for the primary endpoint of 24-hour cough frequency and the secondary endpoint of awake cough frequency was mixed model repeated measures (MMRM) which included the response variable of change from baseline in log-transformed 24-hour cough frequency, adjusted for covariates including treatment, visit, the interaction of treatment by visit, gender, region, the log-transformed baseline value, and the interaction of log-transformed baseline value by visit. An unstructured covariance matrix was used to model the correlation among repeated measurements. Missing data was not imputed in the analysis, assuming missing-at-random (MAR).

The Applicant presents cough frequency results from two analysis methods on the unvalidated original cough data and the validated recount cough data, resulting in a total of 4 analyses: 1) the prespecified MMRM analysis (described by the Applicant as longitudinal ANCOVA) of the unvalidated original cough data, 2) the prespecified MMRM analysis of the validated recount cough data, 3) a post hoc multiple imputation (MI) and ANCOVA method analysis of unvalidated original cough data, 4) a post hoc MI and ANCOVA analysis of validated recount cough data. As stated elsewhere in the BD, we view the recount data as the most reliable dataset because it was generated using a validated algorithm in a blinded manner; thus, the prespecified MMRM model using the validated recount cough data is deemed the primary analysis for efficacy.

For the remainder of secondary endpoints of binary outcomes, which are analyses for treatment comparison to placebo of proportions from baseline at Week 24 (P030) or Week 12 (P027) relative to a threshold (LCQ total score, 24-hour cough frequency, mean weekly CSD total score, and Cough Severity VAS score), the Applicant proposed a logistic regression model, which included covariates for treatment, visit, the interaction of treatment by visit, gender, region, baseline of the underlying continuous response, and the interaction of baseline of the underlying continuous response by visit among subjects who had baseline and post baseline values. Log odds ratio was back transformed into odds ratio for reporting and interpretability. In addition, responder endpoints were analyzed using the Mietinnen and Nurminen method stratified by gender and region.

For analyses of P030 and P027, per the standard regulatory approach to efficacy results, we assessed the results of each individual investigation on its own merits, to assess independent substantiation of results.

3.1.3 Pivotal Trial P030 and P027 Results

Given that the trials were similar in design and utilized the same endpoints, results are presented for each study in parallel within each subsection. The efficacy analyses and discussion will focus on the results for the proposed dosage of gefapixant 45 mg BID dose rather than the 15 mg BID dose.

3.1.3.1 Subject Disposition

Subject disposition at Week 24 in Trial P030 is summarized in Table 5, and disposition at Week 12 in trial P027 is summarized in Table 6. The timepoints correspond to the end of the main study period when the primary efficacy analysis was conducted in each trial. The pivotal trials randomized subjects into one of three study arms using an equal randomization ratio. In both studies, adverse events (AEs), primarily related to taste disturbances, led to dose-dependent treatment discontinuations and study discontinuations. It does not appear that these discontinuation rates impacted the efficacy results meeting statistical significance (see the tipping point analysis on missing data sensitivity for P030 in the Appendix). See Section 6.1.1 for subject disposition at Week 52 for both trials.

Table 5. P030 Subject Disposition Through Week 24, All Subjects as Randomized

Variable	Placebo n (%)	Gefapixant 15 mg BID n (%)	Gefapixant 45 mg BID n (%)	Overall n (%)
Participants in population (N)	436	442	439	1317
Study status				
Completed main study period (Week 24)	382 (87.6)	368 (83.3)	355 (80.9)	1105 (83.9)
Discontinued study	54 (12.4)	74 (16.6)	84 (19.2)	212 (16.1)
Lost to follow-up	6 (1.4)	2 (0.4)	5 (1.1)	13 (1.0)
Physician decision	1 (0.2)	0 (0.0)	3 (0.7)	4 (0.3)
Screen failure	1 (0.2)	2 (0.4)	0 (0.0)	3 (0.2)
Withdrawal by subject	46 (10.6)	68 (15.4)	74 (16.9)	188 (14.3)
Death	0 (0.0)	1 (0.2)	0 (0.0)	1 (0.1)
Other	0 (0.0)	1 (0.2)	2 (0.5)	3 (0.2)
Treatment status for main study period				
Completed treatment at Week 24	369 (84.6)	358 (81.0)	314 (71.5)	1041 (79.0)
Discontinued treatment	66 (15.2)	82 (18.6)	125 (28.5)	273 (20.7)
Adverse event	21 (4.8)	34 (7.7)	88 (20.1)	143 (10.8)
Lost to follow-up	5 (1.1)	2 (0.4)	3 (0.7)	10 (0.8)
Other	1 (0.2)	1 (0.2)	0 (0.0)	2 (0.1)
Physician decision	2 (0.5)	0 (0.0)	2 (0.5)	4 (0.3)
Withdrawal by subject	37 (8.5)	44 (9.9)	30 (6.8)	111 (8.4)
Death	0 (0.0)	1 (0.2)	0 (0.0)	1 (0.1)
Noncompliance with treatment	0 (0.0)	0 (0.0)	2 (0.5)	2 (0.1)

Source: FDA Statistical Analyst; adbase.xpt (Week 24).

Abbreviations: BID, twice daily; N, number of subjects in treatment arm; n, number of subjects in specified population or group

Table 6. P027 Subject Disposition Through Week 12, All Subjects as Randomized

Variable	Placebo n (%)	Gefapixant 15 mg BID n (%)	Gefapixant 45 mg BID n (%)	Overall n (%)
Participants in population (N)	244	244	244	732
Study status				
Completed main study period (Week 12)	199 (81.6)	200 (82.0)	184 (75.4)	583 (79.7)
Discontinued study	45 (18.4)	44 (18.0)	60 (24.5)	149 (20.3)
Death	2 (0.8)	1 (0.4)	0 (0.0)	3 (0.4)
Lost to follow-up	2 (0.8)	1 (0.4)	1 (0.4)	4 (0.5)
Other	1 (0.4)	0 (0.0)	0 (0.0)	1 (0.1)
Physician decision	2 (0.8)	3 (1.2)	3 (1.2)	8 (1.1)
Screen failure	1 (0.4)	0 (0.0)	1 (0.4)	2 (0.3)
Withdrawal by subject	37 (15.2)	39 (16.0)	55 (22.5)	131 (17.9)
Treatment status for main study period				
Completed treatment at Week 12	213 (87.3)	216 (88.5)	182 (74.6)	611 (83.5)
Discontinued treatment	30 (12.3)	28 (11.5)	61 (25.0)	119 (16.3)
Adverse event	7 (2.9)	8 (3.3)	40 (16.4)	55 (7.5)
Death	1 (0.4)	0 (0.0)	0 (0.0)	1 (0.1)
Physician decision	1 (0.4)	0 (0.0)	0 (0.0)	1 (0.1)
Withdrawal by subject	21 (8.6)	20 (8.2)	21 (8.6)	62 (8.5)
Discontinued treatment	30 (12.3)	28 (11.5)	61 (25.0)	119 (16.3)

Source: adbase.xpt; statistical analyst.

Abbreviation: BID, twice daily; N, number of subjects in treatment arm; n, number of subjects in specified population or group

3.1.3.2 Demographic Characteristics

Baseline subject demographic characteristics for P030 and P027 are shown in Table 7. Of the randomized subjects who received at least one dose of study drug, the majority were female, White, and had a mean age of 58 years. Subject demographic characteristics were generally similar between P030 and P027 and between treatment arms within each trial. Overall, the study population appears representative of the general CC patient population, and there were no observed differences in demographics across arms that would be expected to impact the evaluation of efficacy.

Table 7. Demographic Characteristics for Trials P030 and P027 (FAS Population)

	Trial P030			Trial P027		
Characteristic	**Placebo N=435 n (%)**	**Gefapixant 15 mg N=440 n (%)**	**Gefapixant 45 mg N=439 n (%)**	**Placebo N=243 n (%)**	**Gefapixant 15 mg N=244 n (%)**	**Gefapixant 45 mg N=243 n (%)**
Sex (n%)						
Female	326 (74.9)	329 (74.8)	329 (74.9)	181 (74.5)	181 (74.2)	180 (74.1)
Male	109 (25.1)	111 (25.2)	110 (25.1)	62 (25.5)	63 (25.8)	63 (25.9)
Age (years)						
Mean (SD)	58 (12.6)	58.6 (11.4)	57.8 (12.4)	57.9 (13.1)	59.6 (11.7)	59.4 (13.1)
Median (min, max)	60 (19, 84)	60 (22, 88)	59 (19, 87)	61 (21, 81)	61 (22, 89)	61 (19, 85)
Age group, years (n%)						
<45	56 (12.9)	51 (11.6)	63 (14.4)	40 (16.5)	26 (10.7)	31 (12.8)
45-54	88 (20.2)	91 (20.7)	85 (19.4)	42 (17.3)	46 (18.9)	50 (20.6)
55-64	147 (33.8)	155 (35.2)	145 (33.0)	64 (26.3)	80 (32.8)	67 (27.6)
≥65	144 (33.1)	143 (32.5)	146 (33.3)	97 (39.9)	92 (37.7)	95 (39.1)
Race (n%)						
American Indian or Alaska Native	20 (4.6)	28 (6.4)	24 (5.5)	7 (2.9)	6 (2.5)	8 (3.3)
Asian	15 (3.4)	14 (3.2)	15 (3.4)	35 (14.4)	35 (14.3)	34 (14.0)
Black or African American	5 (1.1)	9 (2.0)	14 (3.2)	4 (1.6)	3 (1.2)	4 (1.6)
Multiple	36 (8.3)	31 (7.0)	37 (8.4)	8 (3.3)	5 (2.0)	11 (4.5)
Native Hawaiian or other Pacific Islander	4 (0.9)	2 (0.5)	3 (0.7)	0 (0)	0 (0)	0 (0)
White	355 (81.6)	356 (80.9)	346 (78.8)	189 (77.8)	195 (79.9)	186 (76.5)
Ethnicity (n%)						
Hispanic or Latino	85 (19.5)	93 (21.1)	89 (20.3)	33 (13.6)	35 (14.3)	33 (13.6)
Not Hispanic or Latino	347 (79.8)	345 (78.4)	344 (78.4)	203 (83.5)	205 (84.0)	207 (85.2)
Not reported	2 (0.5)	2 (0.5)	2 (0.5)	5 (2.1)	3 (1.2)	3 (1.2)
Unknown	1 (0.2)	0 (0)	4 (0.9)	2 (0.8)	1 (0.4)	0 (0)
BMI (kg/m^2)						
Mean (SD)	28.6 (5.6)	28.9 (6.1)	28.9 (5.7)	28 (5.6)	28.7 (5.8)	28 (5.9)
Median (min, max)	27.5 (17.1, 53)	28 (17.3, 56.3)	28.2 (16.5, 53.1)	27.4 (17.6, 52.6)	28 (16.7, 50.3)	26.9 (15.8, 53.6)
Region (n%)						
Asia Pacific	26 (6.0)	27 (6.1)	28 (6.4)	35 (14.4)	34 (13.9)	34 (14.0)
Europe	238 (54.7)	238 (54.1)	239 (54.4)	121 (49.8)	123 (50.4)	121 (49.8)
North America	97 (22.3)	99 (22.5)	98 (22.3)	56 (23.0)	55 (22.5)	56 (23.0)
Others	74 (17.0)	76 (17.3)	74 (16.9)	31 (12.8)	32 (13.1)	32 (13.2)

Source: FDA clinical data scientist.

Data reflect planned allocation at baseline regardless of actual treatment received.

* FAS differs from all subjects as randomized because subjects who did not receive treatment were not included in the FAS

Abbreviations: BMI, body mass index; FAS, full analysis set, N, number of subjects in treatment group; n, number of subjects with given characteristic; SD, standard deviation

3.1.3.3 Baseline Disease Characteristics

Subject baseline disease characteristics were generally similar between P030 and P027 and between treatment arms within each trial. These data are shown in Table 8. As noted above, there were no eligibility criteria for baseline cough frequency. For all treatment arms, the mean baseline cough frequency was between 25 to 30 cough frequency, with the exception of the placebo arm in P027, which had a mean baseline cough count of 38 cough frequency due to an outlier of 1053 cough frequency. Notably, there was a large standard deviation for mean baseline cough frequency in all arms. The variability in baseline cough frequency is unlikely to have impacted the statistical analysis of the primary endpoint given the analysis approach, though it may create challenges for clinical application of the trial results to an individual CC patient. Other baseline disease characteristics are generally well-matched. Overall, baseline disease characteristic differences across arms are not likely to have impacted the efficacy results.

Table 8. Baseline Disease Characteristics for Trials P030 and P027, FAS Population

	Trial P030			Trial P027		
Characteristic	**Placebo N=435 n (%)**	**Gefapixant 15 mg N=440 n (%)**	**Gefapixant 45 mg N=439 n (%)**	**Placebo N=243 n (%)**	**Gefapixant 15 mg N=244 n (%)**	**Gefapixant 45 mg N=243 n (%)**
Primary diagnosis (n, %)						
Refractory CC	278 (63.9)	273 (62.0)	279 (63.6)	148 (60.9)	141 (57.8)	139 (57.2)
Unexplained CC	157 (36.1)	167 (38.0)	160 (36.4)	95 (39.1)	103 (42.2)	104 (42.8)
Duration of CC (n, %)						
<10 years	247 (56.8)	231 (52.5)	258 (58.8)	127 (52.3)	130 (53.3)	134 (55.1)
≥10 years	188 (43.2)	209 (47.5)	181 (41.2)	116 (47.7)	114 (46.7)	109 (44.9)
Duration of CC (years)						
Mean (SD)	10.7 (8.8)	11.9 (10.7)	10.9 (9.9)	11.7 (9.9)	11.8 (9.1)	11.2 (9.4)
Median (min, max)	8 (2, 51)	9 (1, 75)	7 (2, 65)	9 (2, 59)	9 (2, 45)	8 (2, 56)
Mean weekly Cough Severity VAS (n, %)						
<60 mm	135 (31.0)	149 (33.9)	128 (29.2)	62 (25.5)	77 (31.6)	71 (29.2)
≥60 mm	299 (68.7)	290 (65.9)	309 (70.4)	179 (73.7)	167 (68.4)	172 (70.8)
Missing	1 (0.2)	1 (0.2)	2 (0.5)	2 (0.8)	0 (0)	0 (0)
Mean weekly cough severity VAS (mm)						
Mean (SD)	68.5 (14.3)	67.4 (14.8)	67.7 (13.9)	69.1 (13.9)	68.2 (15)	67.9 (12.8)
Median (min, max)	68.1 (31.4, 100)	66.6 (20.7, 100)	67 (25.6, 99.3)	69 (37.6, 100)	67.5 (30.4, 100)	67.6 (27, 99.3)
24-Hour cough frequency (n, %)						
<20 coughs/hour	210 (48.3)	201 (45.7)	227 (51.7)	96 (39.5)	114 (46.7)	121 (49.8)
≥20 coughs/hour	222 (51.0)	230 (52.3)	207 (47.2)	136 (56.0)	121 (49.6)	116 (47.7)
Missing	3 (0.7)	9 (2.0)	5 (1.1)	11 (4.5)	9 (3.7)	6 (2.5)

Characteristic	Trial P030 Placebo N=435 n (%)	Trial P030 Gefapixant 15 mg N=440 n (%)	Trial P030 Gefapixant 45 mg N=439 n (%)	Trial P027 Placebo N=243 n (%)	Trial P027 Gefapixant 15 mg N=244 n (%)	Trial P027 Gefapixant 45 mg N=243 n (%)
24-hour cough frequency (coughs/hour)						
Mean (SD)	28.5 (24.6)	28.1 (22.2)	28.6 (29.9)	39.5 (81.1)	28.0 (22.0)	30.2 (39.4)
Median (min, max)	21.3 (0.7, 183.6)	22.1 (1.0, 151.6)	19.9 (0.2, 230.1)	26.1 (0.3, 1053.5)	21.8 (0.8, 132.8)	20.9 (0.2, 399.1)
Q1, Q3	12.4, 37.1	11.4, 39.5	10.9, 37.9	12.9, 45.5	13.2, 37.3	12.2, 36.2
Geometric mean[1]	20.4	20.2	19.4	23.6	20.9	18.9

Source: adsl.xpt, adbase.xpt; software: R.

Data reflect planned allocation at baseline regardless of actual treatment received.

[1] Geometric mean was calculated by FDA statistical reviewer using SAS.

Abbreviations: CC, chronic cough; FAS, full analysis set; N, number of subjects in treatment group; n, number of subjects with given characteristic; Q, quartile; SD, standard deviation; VAS, visual analog scale

3.1.3.4 Primary Endpoint Results

As discussed above, the FDA conducted all analyses of cough frequency on the validated cough *recount* data as it was produced via a validated process in a blinded manner, ensuring reliability and accuracy of the efficacy results. As described in the SAPs, the prespecified primary analysis for 24-hour cough frequency was to use an MMRM analysis (referred to by the Applicant as longitudinal ANCOVA). Analysis results for the primary endpoint of 24-hour cough frequency based on the recount data for Trials P030 and P027 using the MMRM model are presented in Table 9. For P030, the relative reduction in geometric mean ratio of 24-hour cough frequency at Week 24 from baseline was 14.6% (95% CI, -26.0 to -1.5, p-value: 0.030) with gefapixant 45 mg compared to placebo. Analysis results for P027 yielded a 17.0% relative reduction (95% CI, -31.5 to 0.6, p-value: 0.057). Results at the 15 mg dose were not statistically significant in either study. Several sensitivity analyses were conducted on the primary endpoint to assess the robustness of the primary analysis results and found a consistent magnitude of treatment effect (see details in Section 6.1.6).

Table 9. Analysis of 24-Hour Cough Frequency, Trials P030 and P027 (Full Analysis Set, Recount Data, MMRM)

Statistic	Trial P030 (Week 24) Placebo	Trial P030 (Week 24) Gefapixant 15 mg	Trial P030 (Week 24) Gefapixant 45 mg	Trial P027 (Week 12) Placebo	Trial P027 (Week 12) Gefapixant 15 mg	Trial P027 (Week 12) Gefapixant 45 mg
N	419	415	409	222	227	217
Geometric mean at baseline	20.4	20.2	19.4	23.6	20.9	18.9
Geometric mean at Week 24 or 12	8.7	8.3	7.1	10.6	10.2	7.4
Geometric mean ratio*	0.43	0.41	0.37	0.45	0.49	0.39
Model based geometric mean ratio (95% CI)**	0.43 (0.39, 0.48)	0.42 (0.38, 0.47)	0.37 (0.33, 0.41)	0.47 (0.41, 0.54)	0.48 (0.42, 0.56)	0.39 (0.34, 0.45)
Relative reduction (%) in geometric mean ratio (95% CI)***		-2.7 (-15.6, 12.1)	-14.6 (-26.0, -1.5)		3.1 (-14.5, 24.4)	-17.0 (-31.5, 0.6)
p-value		0.703	0.030		0.748	0.057

Source: adeff.xpt; Table 4-2 in CSR Addenda for P030 and P027, validated by statistical analyst.

Geometric means were used because these frequency data were log-transformed.

* Based on subjects with non-missing values at baseline and Week 24 (P030) and Week 12 (P027).

** Based on the MMRM model.

*** The estimated relative reduction (relative to placebo) is calculated by 100 (e^{DIFF} -1). DIFF is the treatment difference in change from baseline at Week 24 (P030) or Week 12 (P027) based on the log-transformed data.

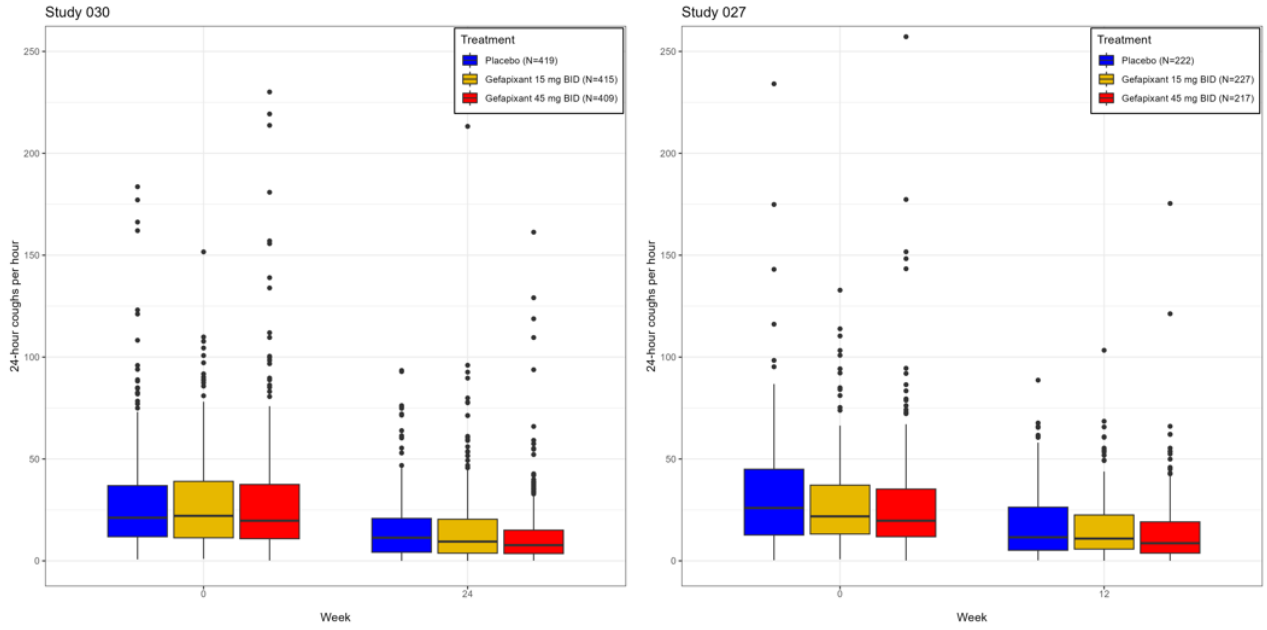
Abbreviations: CI, confidence interval; CSR, clinical study report; MMRM, mixed model repeated measures; N, number of subjects who had baseline and postbaseline assessments

The primary endpoint was analyzed as the mean change from baseline in the natural log-transformed cough frequency at Week 24 or 12, in Trials P030 and P027 respectively, and characterized as the relative reduction over placebo. It is important to note the large placebo effect in both trials. The prespecified analysis of the recount cough data yielded a statistically significant difference between geometric mean ratios of the proposed 45 mg dose and placebo in Trial P030 only. We note that use of the recount data shifted the p-value to >0.05 for the smaller Trial P027, but the point estimate for the reduction in cough frequency was similar in both trials, a relative reduction in the geometric mean ratio of -15% to -17% compared to placebo from baseline to Week 24 or Week 12. The similarity in the small treatment differences and the shift in statistical significance may indicate a lack of robustness in the treatment effect. However, because achievement of statistical significance alone does not confirm that the level of response corresponds to a clinically meaningful benefit to patients, the central question for the Committee's input remains whether there is sufficient evidence to determine whether these results are clinically meaningful to patients.

Because features of the primary endpoint analysis (i.e., log-transformed data, relative reduction versus placebo in geometric mean ratio) make clinical interpretation of the results challenging, FDA conducted descriptive graphical analyses to assist in interpreting the magnitude of the treatment effect. We conducted post hoc analyses of the absolute cough frequency, a more intuitive expression of the primary endpoint, which revealed small differences between treatment groups in the median cough frequency. The median change in 24-hour cough frequency from baseline was -9.8 and -10.5 in the gefapixant 45 mg groups and -8.7 and -8.9 in the placebo groups at Week 24 (P030) and Week 12 (P027), respectively (Table 25 and Table 26).

Figure 4 is a boxplot for each trial showing the 24-hour coughs frequency at baseline and at Week 24 for Trial P030 and at Week 12 for Trial P027. These displays allow visualization of aspects of the cough frequency data that may be useful for clinicians, including the median cough frequencies at baseline and Week 12 or 24, the magnitude of cough frequency reduction over the treatment period, and the similarity of cough frequencies across the three treatment arms.

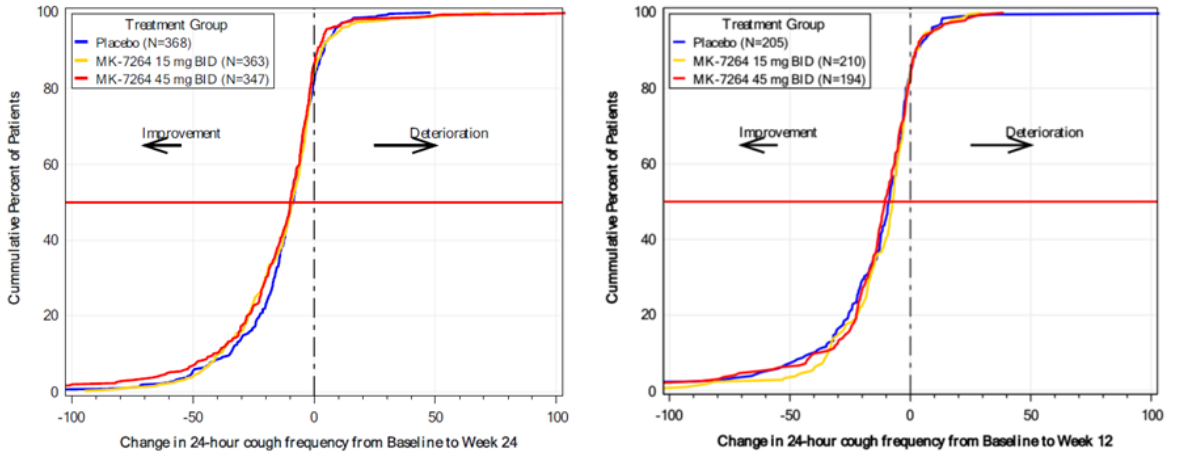
Figure 4. Twenty-Four-Hour Cough Frequency for Trials P030 and P027 (Y-Axis Restricted to 250 Coughs per Hour, Full Analysis Set*, Recount Data)



Source: Statistical analyst.
The upper/lower bounds of the boxes and the horizontal line in each box indicate 25th/75th percentiles and the median, respectively.
Observed data at Week 12/24 were displayed.
For Trial P027, two subjects in the gefapixant 45 mg group at baseline, four subjects in the placebo group at baseline, and two subjects in the placebo group at week 12 had 24-hour cough frequency values greater than 250 and were not included in the figure. These eight subjects had 24-hour cough frequency ranged from 257.2 to 1053.5.
* Full analysis set, with outliers above 250 cough frequency excluded, for purposes of examining the majority of data in the trials
Abbreviation: N, number of subjects who had baseline and postbaseline assessments

Figure 5 shows the empirical cumulative distribution function plots of change from baseline in 24-hour cough frequency at Week 24 for P030 and at Week 12 for P027, respectively. The observed treatment effect was small, as evidenced by the minimal separation between all treatment arms in both studies.

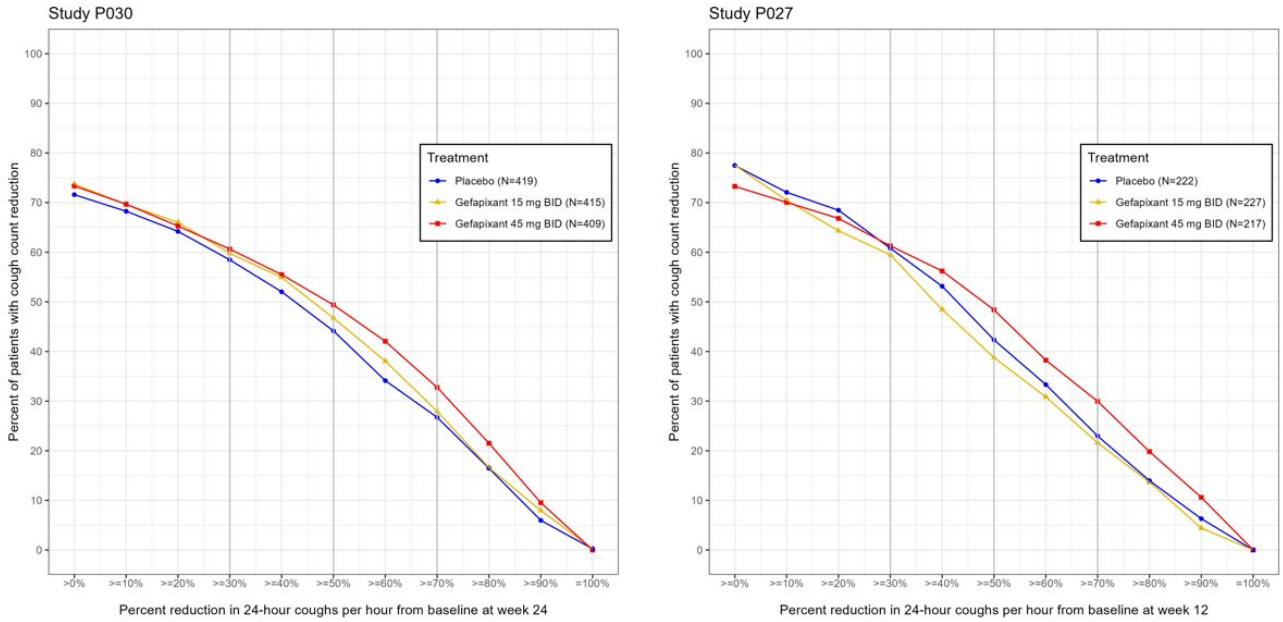
Figure 5. Empirical Cumulative Distribution Function Plot of Change From Baseline in 24-Hour Cough Frequency at Week 24 (Trial P030) and at Week 12 (Trial P027)



Source: Applicant's Figures 2-3 and 1b-1 in response to FDA's September 19, 2023, information request.
Abbreviations: BID, twice daily; FDA, Food and Drug Administration

Figure 6 shows the cumulative responder curves by varying responder thresholds for percent reduction from baseline in 24-hour cough frequency to Week 24 for Trial P030 and to Week 12 for Trial P027, respectively. The vertical lines at 30%, 50%, and 70% reductions from baseline in 24-hour cough frequency display the Applicant's pre-specified thresholds of responses for analysis as secondary and exploratory endpoints along the continuum of response from 0 to 100%. As shown in Figure 6, there is a large proportion of placebo responders that tracks with the gefapixant responders. While there is a numerical difference in the proportion of responders for percent reduction in cough frequency between gefapixant 45mg and placebo, the magnitude of the difference is quite small. For example, in P027, ≥50% reduction in cough frequency was observed in 48% of subjects who received gefapixant 45 mg vs. 42% subjects who received placebo. In other words, the proportion of subjects who had ≥50% reduction in cough frequency was only 6% more in the gefapixant 45 mg group compared to placebo.

Figure 6. Proportion of Subjects With Percentage Reduction From Baseline in 24-Hour Cough Frequency at Week 24 (Trial P030) and Week 12 (Trial P027) (FAS, Recount Data)

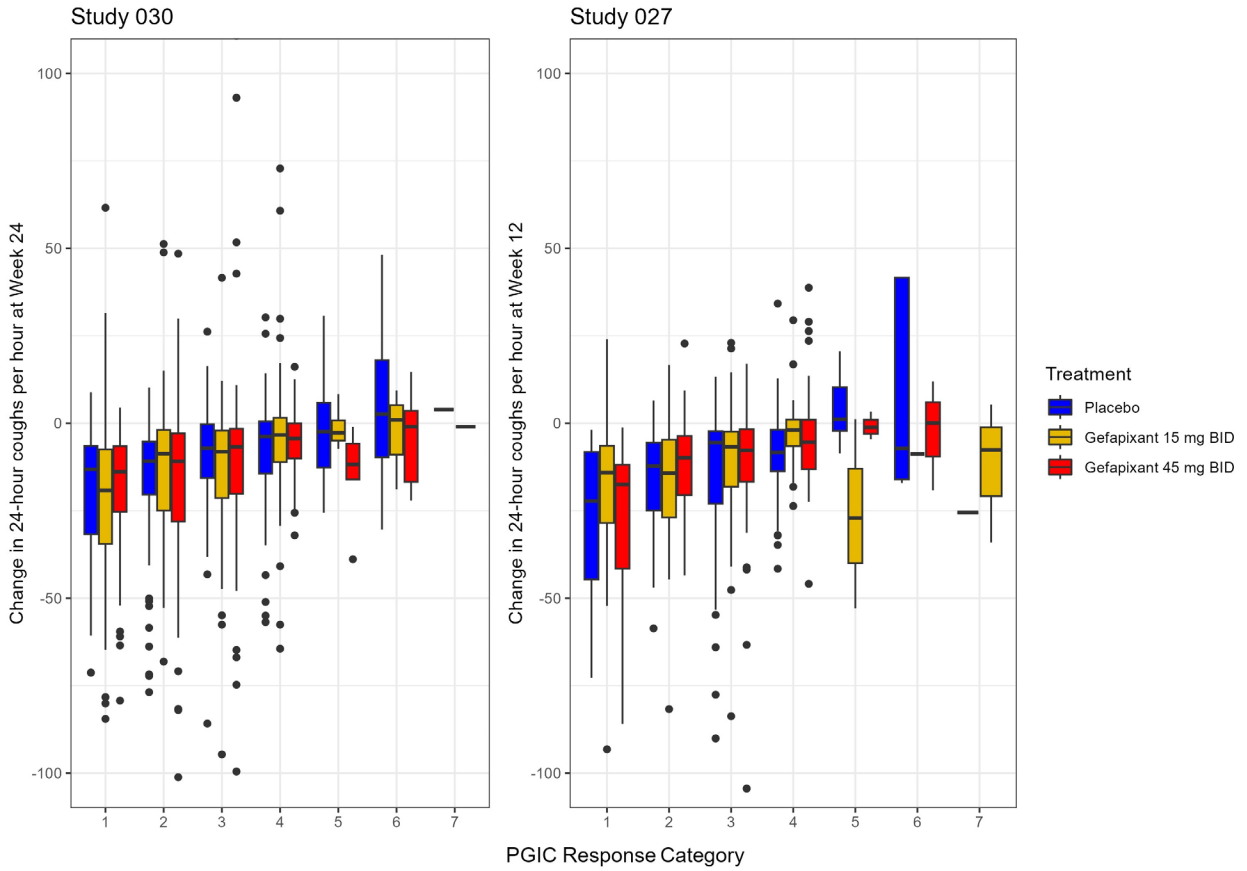


Source: Statistical analyst. Refer to Table 32 and Table 33.
For percentage reduction, denominator was number of subjects who had baseline and postbaseline assessments, and numerator was number of subjects whose percent change in 24-hour cough frequency reached individual thresholds. Subjects who had missing data at the end of the main study period were treated as nonresponders.
Abbreviation: BID, twice daily; FAS, full analysis set

To probe the clinical meaningfulness of the primary endpoint results, the FDA requested post hoc anchor-based analyses using PGIC[4]. The PGIC is a single-item PRO asking patients to describe their cough "now" compared to the start of treatment with seven response options ranging from "very much improved" to "very much worse." Using the PGIC as an anchor, we can explore whether there is a correlation between patient-reported improvement in cough and the change in cough frequency. The Applicant reports numerical differences in "improvement responders" between treatment groups in both P027 and P030; as with other efficacy outcomes, there was a large placebo response. However, both trials showed low correlation between the PGIC and change in cough frequency (Polyserial/Spearman: 0.15/0.32 and 0.23/0.30 for Trials P027 and P030, respectively). The poor correlation is evidenced by the boxplot depicted in Figure 7, which indicates that patient-reported improvement in chronic cough (as captured by PGIC) did not track with degree of reduction in cough frequency. In other words, patients who reported feeling better per the PGIC were not necessarily those patients who were coughing less.

[4] The PGIC is the only PRO measure administered in both studies that would be considered reasonable as an anchor scale.

Figure 7. Change From Baseline in 24-Hour Cough Frequency at Week 24 (Trial P030) and at Week 12 (Trial P027) by Response Category on PGIC (Y-Axis Restricted to ±100 Coughs per Hour, Full Analysis Set, Recount Data)



Source: Statistical analyst. Refer to Table 35 (P027) and Table 34 (P030)
PGIC response category: 1="Very much improved", 2="Much improved", 3="Minimally improved", 4="No change", 5="Minimally worse", 6="Much worse", 7="Very much worse."
Abbreviations: BID, twice daily; PGIC, patient global impression of change

Additionally, taste-related disturbances with gefapixant use occurred in about 65% of subjects in the gefapixant 45 mg BID group compared to 7% in the placebo group. All subjects and investigators were aware of the frequent occurrence of taste disturbance, which was disclosed in the informed consent and investigator brochure. Knowledge of this likely pharmacologic effect may have introduced bias to the PRO results, based on awareness of assigned treatment, in ways that are difficult to predict and quantify. FDA performed post hoc, exploratory analyses of the relationship of taste-related adverse event (AE) incidence and the primary endpoint results (see Section 6.1.7), which suggest a potential association between incidence of taste-related AEs and reduction in cough frequency. Conclusions cannot be made regarding the true impact of taste disturbance on study outcomes, yet these concerns introduce uncertainty into the efficacy results.

3.1.3.5 Secondary Endpoint Results (Multiplicity Controlled)

The multiplicity control strategy for endpoints in Trials P030 and P027 are described in Table 4. In Table 10 and Table 11, the results for secondary endpoints for P030 and P027 are shown according to their placement within the multiplicity control hierarchy; this is followed by a discussion of the results for each endpoint. The remaining comparisons evaluated gefapixant at the lower dose of 15 mg BID; the

results are shown for completeness but not discussed in detail as the Applicant is not seeking approval for the lower dose and there was no statistically significant difference between the lower dose and placebo on the primary endpoint.

Table 10. Trial P030 Secondary Endpoints in the Multiplicity Testing Hierarchy (Full Analysis Set, Recount Data)

Endpoint	Statistic	Placebo	Gefapixant 15 mg	Gefapixant 45 mg
Awake cough frequency	N[a]	419	415	409
	Geometric mean at baseline	26.9	26.6	25.2
	Geometric mean at Week 24	11.3	10.6	9.0
	Geometric mean ratio (Week 24÷baseline) (95% CI)[b c]	0.43 (0.38, 0.47)	0.41 (0.36, 0.45)	0.36 (0.32, 0.40)
	Relative reduction (%) (95% CI)[c]		-4.6 (-17.4, 10.2)	-15.5 (-27.0, -2.3)
	p-value		0.521	0.023
≥1.3-point increase from baseline in LCQ total score at Week 24	N[d]	355	352	342
	n[d] (%)	245 (69)	264 (75)	262 (77)
	Odds ratio vs. placebo[e] (95% CI)		1.3 (1.0, 1.8)	1.4 (1.0, 2.0)
	p-value		0.077	0.040
	Responder (%)[f]	245/415 (59)	264/417 (63)	262/419 (63)
	Estimated difference (%) (95% CI)[g]		4.2 (-2.4, 10.8)	3.3 (-3.3, 9.9)
≥30% reduction from baseline in 24-hour cough frequency at Week 24	N[d]	368	363	347
	n[d] (%)	245 (67)	248 (68)	248 (72)
	Odds ratio vs. placebo[e] (95% CI)		1.0 (0.8, 1.4)	1.2 (0.9, 1.7)
	p-value		0.905	0.188
	Responder (%)[f]	245/432 (57)	248/431 (58)	248/434 (57)
	Estimated difference (%) (95% CI)[g]		0.9 (-5.8, 7.4)	0.4 (-6.1, 7.0)

Source: adeff.xpt; Tables 4-4 and 4-10 in CSR Addenda, Table 14.2-22 in CSR of Trial P030, validated by the statistical review team.

[a] N=Number of subjects who had baseline and postbaseline assessments.

[b] Based on subjects with nonmissing values at baseline and Week 24.

[c] Based on the MMRM model. The estimated relative reduction (relative to placebo) is calculated by 100 (e^{DIFF} -1). DIFF is the treatment difference in change from baseline at Week 24 based on the log-transformed data.

[d] N=Number of subjects with available data at Week 24; n=number of responders at Week 24.

[e] Based on the logistic regression model with terms of treatment, visit, treatment-by-visit interaction, gender, region, baseline, and baseline-by-visit interaction among subjects who had baseline and post baseline values.

[f] N=number of subjects who had baseline values. Subjects with nonmissing baseline data but missing Week 24 data are imputed as nonresponders.

[g] Based on the Mietinnen and Nurminen method.

Abbreviations: CI, confidence interval; CSR, clinical study report; Leicester Cough Questionnaire (LCQ)

Table 11. Trial P027 Secondary Endpoints in the Multiplicity Testing Hierarchy (Full Analysis Set, Recount Data)

Endpoint	Statistic	Placebo	Gefapixant 15 mg	Gefapixant 45 mg
Awake cough frequency	N[a]	222	227	217
	Geometric mean at baseline	31.4	217.1	25.0
	Geometric mean at Week 12	13.8	13.2	9.6
	Geometric mean ratio (Week 12/baseline) (95% CI)[b,c]	0.46 (0.40, 0.53)	0.48 (0.41, 0.55)	0.39 (0.33, 0.45)
	Relative reduction (%) (95% CI)[c]		4.0 (-14.2, 26.0)	-16.3 (-31.2, 1.9)
	p-value		0.691	0.076
≥30% reduction from baseline in 24-hour cough frequency at Week 12	N[d]	205	210	194
	n[d] (%)	135 (66)	135 (64)	133 (69)
	Odds ratio vs. placebo[e] (95% CI)		1.0 (0.7, 1.6)	1.2 (0.8, 1.8)
	p-value		0.899	0.435
	Responder (%)[f]	135/232 (58)	135/235 (57)	133/237 (56)
	Estimated difference (%) (95% CI)[g]		-0.9 (-9.9, 8.0)	-2.2 (-11.2, 6.8)

Source: adeff.xpt; Tables 4-3, 4-4 in CSR Addenda of Trial P027, validated by the statistical review team.
Footnotes for Trial P027 are identical to those for Trial P030, except the timepoint is Week 12.
Abbreviations: CI, confidence interval; CSR, clinical study report

3.1.3.5.1 Awake Cough Frequency

This secondary endpoint is intended to capture the frequency of coughs while the subject is awake during a 24-hour period, excluding nighttime sleep. The VitaloJAK cough analyst determined awake and sleep states using a protocolized method that incorporated the local time of the recording and decreased audio activity in the recording. Similar to the primary endpoint results, awake cough frequency was statistically significantly reduced in the gefapixant 45 mg group compared to placebo in Trial P030 only; however, the estimated relative reduction (in geometric mean ratio) was similar across trials (-15.5% and -16.3% in P030 and P027, respectively; see Table 11).

Awake cough frequency is conceptually quite similar to the primary endpoint (cough frequency over a 24-hour period, whether or not the subject was awake); consequently, the results are similar to the primary endpoint. Given that awake cough frequency is a subset of the 24-hour cough frequency primary endpoint, the same questions raised around the clinical meaning of the numerically small reductions in cough frequency compared to placebo apply to this endpoint as well.

3.1.3.5.2 LCQ Total Score

The LCQ (see Copy of the LCQ) is a 19-item PRO instrument that assesses cough symptoms and its impacts over a 2-week recall period using a 7-point Likert scale ranging from 1 to 7. The LCQ is comprised of three domains: physical, psychological, and social (see Conceptual Framework of the LCQ). The sum of individual domain scores, which range from 1 to 7, provides the total LCQ score, ranging from 3 to 21. Higher scores indicate better health status. FDA has concerns about whether the LCQ total score is fit-for-purpose. Specifically, concepts evaluated in the social and psychological domains (e.g., embarrassed or worried about cough; cough has interfered with enjoyment of life) can be influenced by factors outside of treatment. In addition, there is a lack of information to determine what change in score is clinically meaningful, and the Applicant did not provide sufficient evidence to support a responder threshold.

For Trial P030 only, the Applicant prespecified a multiplicity-controlled responder analysis of the LCQ total score using a threshold of ≥1.3 points to define responders. At Week 24, the proportion of subjects

with a ≥1.3-point increase from baseline in LCQ total score was statistically significantly higher in the gefapixant 45 mg group compared with the placebo group (Table 10; odds ratio, 1.4 (95% CI [1.0, 2.0]). While the OR for the LCQ total score in P030 was statistically significant, the proportion of patients who had a ≥1.3-point increase from baseline in the LCQ total score is only 3.3% more than placebo (Table 10). This numerically small difference in responders between gefapixant and placebo is of questionable clinical significance. In addition, we question whether a change as small as 1.3 points on the LCQ total score, which has a maximum possible score of 21, is inherently meaningful to patients. Further, the magnitude of the effect on the LCQ total score was nearly the same in both dosing arms receiving gefapixant, despite the 15 mg dose having no detectable impact on cough frequency endpoints compared to placebo (see Table 10). We also note that over two-thirds of subjects met the definition of a responder regardless of whether they received gefapixant or placebo.

To assist with interpretation of the responder analysis on LCQ total score, FDA conducted post hoc analyses of the change from baseline in total score and in each of the individual domains contributing to the LCQ total score. The absolute change from baseline in the LCQ total score compared to placebo was small (<1 point) in relationship to the range of possible scores (0.78 points at Week 24 and 0.35 points at Week 12 in P030 and P027, respectively, out of a possible range of 3 to 21).

Despite the statistical significance of this secondary endpoint, it is challenging to determine what this responder analysis result means clinically to patients. Additional post hoc analyses on the LCQ physical domain score are described in Section 3.1.3.6.

3.1.3.5.3 Cough Count Reduction ≥30%

For both trials, the Applicant prespecified a responder analysis for ≥30% reduction from baseline in cough frequency at the end of the main study period. The gefapixant 45 mg groups had a small increase in the proportion of "responders" compared to placebo (at Week 24 in P030, Table 10, and at Week 12 in P027, Table 11); however, the difference was not statistically significant in either study. We note that a ≥30% reduction was achieved by most subjects regardless of assigned treatment group (66% to 67% of placebo compared to 69% to 72% of gefapixant 45 mg), a finding which calls into question the clinical meaningfulness of a 30% reduction threshold. Refer to Section 3.1.3.4 for additional discussion on other responder cutoffs as shown in Figure 6.

3.1.3.5.4 Secondary Endpoints (not Multiplicity Controlled)

The remaining prespecified secondary endpoints for P030 and P027 were not included in the multiplicity control hierarchy. Given that this is a novel therapeutic area, the FDA reviewed these data as exploratory analyses to further our understanding of the PROs captured in the gefapixant development program. Upon review, FDA has identified limitations and uncertainties with the CSD total score and Cough Severity VAS (discussed below) and the responder threshold cutoff(s) selected for each. Because of these concerns and the lack of multiplicity control in testing for statistical significance, the results should be interpreted with caution.

Analyses of these secondary endpoint results at Weeks 24 and 12 for Trials P030 and P027, respectively, are shown in Table 12, followed by a discussion of the PRO-based endpoints below.

Table 12. Prespecified Secondary Endpoint Analyses Not in the Multiplicity Hierarchy (FAS)

Secondary Endpoint	Statistic	Placebo	Gefapixant 15 mg BID	Gefapixant 45 mg BID
Trial P030, Week 24				
≥1.3-point reduction from baseline in mean weekly CSD total score at Week 24	N[a]	346	336	331
	n[a] (%)	237 (68)	253 (75)	253 (76)
	Odds ratio vs. placebo[b] (95% CI)		1.3 (1.0, 1.8)	1.5 (1.1, 2.1)
	Responder (%)[c]	237/434 (55)	253/439 (58)	253/437 (58)
	Estimated difference (%) (95% CI)[d]		3.1 (-3.5, 9.6)	3.2 (-3.4, 9.8)
≥2.7-point reduction from baseline in mean weekly CSD total score at Week 24	N[a]	346	336	331
	n[a] (%)	154 (45)	162 (48)	186 (56)
	Odds ratio vs. placebo[b] (95% CI)		1.3 (0.9, 1.7)	1.8 (1.3, 2.4)
	Responder (%)[c]	154/434 (36)	162/439 (37)	186/437 (43)
	Estimated difference (%) (95% CI)[d]		1.5 (-4.8, 7.9)	7.1 (0.6, 13.5)
≥30 mm reduction from baseline in Cough Severity VAS score at Week 24	N[a]	346	336	331
	n[a] (%)	150 (43)	178 (53)	178 (54)
	Odds ratio vs. placebo[b] (95% CI)		1.5 (1.1, 2.1)	1.7 (1.2, 2.2)
	Responder (%)[c]	150/434 (35)	178/439 (41)	178/437 (41)
	Estimated difference (%) (95% CI)[d]		6.1 (-0.4, 12.4)	6.2 (-0.3, 12.6)
Trial P027, Week 12				
≥1.3 increase from baseline in LCQ total score at Week 12	N[a]	196	200	194
	n[a] (%)	123 (63)	139 (70)	134 (69)
	Odds ratio vs. placebo[b] (95% CI)		1.4 (0.9, 2.1)	1.3 (0.9, 2.0)
	Responder (%)[c]	123/229 (54)	139/233 (60)	134/236 (57)
	Estimated difference (%) (95% CI)[d]		5.9 (-3.1, 14.9)	2.8 (-6.2, 11.7)
≥1.3 reduction from baseline in mean weekly CSD total score at Week 12	N[a]	211	217	204
	n[a] (%)	112 (53)	137 (63)	129 (63)
	Odds ratio vs. placebo b (95% CI)		1.5 (1.0, 2.2)	1.4 (0.9, 2.1)
	Responder (%)[c]	112/241 (47)	137/244 (56)	129/243 (53)
	Estimated difference (%) (95% CI)[d]		9.9 (1.0, 18.5)	6.5 (-2.3, 15.3)
≥2.7 reduction from baseline in mean weekly CSD total score at Week 12	N[a]	211	217	204
	n[a] (%)	65 (31)	87 (40)	84 (41)
	Odds ratio vs. placebo[b] (95% CI)		1.5 (1.0, 2.3)	1.7 (1.1, 2.5)
	Responder (%)[c]	65/241 (27)	87/244 (36)	84/243 (35)
	Estimated difference (%) (95% CI)[d]		8.7 (0.5, 16.9)	7.4 (-0.8, 15.6)
≥30 mm reduction from baseline in Cough Severity VAS score at Week 12	N[a]	211	217	204
	n[a] (%)	63 (30)	79 (36)	87 (43)
	Odds ratio vs. placebo[b] (95% CI)		1.3 (0.9, 1.9)	1.5 (1.0, 2.3)
	Responder (%)[c]	63/241 (26)	79/244 (32)	87/243 (36)
	Estimated difference (%) (95% CI)[d]		6.4 (-1.8, 14.5)	9.4 (1.2, 17.7)

Source: adeff.xpt; Tables 11-5, 11-6, 11-7, 14.2-35, 14.2-36, 14.2-43 in CSR of Trial P030; Tables 11-7, 11-4, 11-5, 11-6, 14.2-52, 14.2-34, 14.2-37, 14.2-44 in CSR of Trial P027, validated by statistical review team. Figure presentation is Figure 15 and Figure 16.

[a] N=Number of subjects with available data at Week 24; n=number of responders at Week 24.

[b] Based on the logistic regression model with terms of treatment, visit, treatment-by-visit interaction, gender, region, baseline, and baseline-by-visit interaction among subjects who had baseline and post baseline values.

[c] N=number of subjects who had baseline value. Subjects with nonmissing baseline data but missing Week 24 for Trial P030 (Week 12 for Trial P027) data are imputed as nonresponders.

[d] Based on the Miettinen and Nurminen method.

Abbreviations: CI, confidence interval; CSD, Cough Severity Diary; FAS, full analysis set; LCQ, Leicester Cough Questionnaire; VAS, visual analog scale

CSD Total Score

The CSD (see Copy of the CSD) is a 7-item, disease-specific PRO instrument completed daily and intended to assess the frequency (3 items), intensity (2 items), and disruptiveness (2 items) of cough. Each item is rated on a 0 to 10 numeric rating scale with higher scores indicating greater severity. The total daily CSD score was the sum of these 7-item scores, and the mean total daily score was the 7-item sum divided by 7. The mean weekly CSD total score was defined as the average of the mean total daily scores for a given week (range 0 to 10). The baseline mean weekly CSD score was calculated for the week preceding the first dose of study intervention.

For both trials, the Applicant prespecified a responder analysis of reduction from baseline in weekly CSD total score at the end of the main study period, selecting 2 different thresholds of ≥1.3 and ≥2.7-point reduction to define responders. During development, FDA requested that the Applicant provide evidence to demonstrate that the responder thresholds reflect meaningful change from the patient perspective; the Applicant has not addressed these concerns. The odds ratios in the gefapixant 45 mg versus placebo arm were nominally significant with OR of 1.4 to 1.5 for the lower threshold (≥1.3-point reduction) and OR of 1.7 to 1.8 for the higher threshold (≥2.7-point reduction). Despite their nominal significance, the results for the 45 mg versus placebo comparison need to be interpreted with caution because there is no evidence to support that the proposed responder thresholds represent meaningful within-patient change. FDA's concerns with the CSD and related endpoints are discussed further in Section 6.4.3.

Cough Severity VAS

The Cough Severity VAS (see Copy of the Cough Severity VAS) is a single-item PRO instrument completed each evening, asking the participant to rate the severity of their cough "today" using a 100 mm VAS with "No Cough" at 0 and "Extremely Severe Cough" at 100. The mean weekly VAS score was defined as the average of the mean total daily scores for a given week (range 0 to 100). The baseline mean weekly VAS score was calculated for the week preceding the first dose of study intervention.

For both trials, the Applicant prespecified a responder analysis of reduction from baseline in Cough Severity VAS score at the end of the main study period, proposing a threshold of ≥30 mm reduction to define responders. The odds ratio of 1.5 to 1.7 in the gefapixant 45 mg versus placebo arm was nominally significant. Despite its nominal significance, the result for the 45 mg versus placebo comparison needs to be interpreted with caution because the Applicant has not provided sufficient justification that the proposed cutoff of ≥30 mm reduction represents a meaningful within-patient change. FDA's concerns with the Cough Severity VAS and related endpoints are discussed further in Section 6.4.4.

3.1.3.6 Post hoc Exploratory Analysis – LCQ Physical Domain Score

As described in Section 3.1.3.5.2, FDA has concerns about the LCQ total score, namely, the inclusion of concepts that can be influenced by factors not directly related to treatment intervention and the lack of evidence to support responder thresholds as clinically meaningful. To assist with the interpretation of the responder analysis on LCQ total score, the FDA conducted post hoc, exploratory analyses of each of the individual domains contributing to the LCQ total score. The physical domain is of particular interest because it more directly assesses cough and its impact, and therefore, may be more relevant to informing regulatory decisions than the psychological and social domains. The LCQ physical domain (see Section 6.4.5.1) comprises eight items capturing cough-related symptoms such as chest or stomach

pains, sputum (phlegm) production, being tired, cough on exposure to paints or fumes, sleep disturbance, coughing bouts, hoarse voice, and having energy. The physical domain score is calculated by summing item scores divided by number of items (n=8) included in this domain. The range of possible scores for the LCQ physical domain is 1 to 7, with a higher score indicating better health status. Longitudinal assessment of the LCQ physical domain score over the treatment course may be a useful tool to understand how patients feel or function because of the condition and its treatment.

The LCQ physical domain appears to contribute slightly less towards the improvement in total score than the other, less relevant domains (data not shown). Further, the post hoc analysis on change from baseline in the LCQ physical domain score at Week 12 (for P027) and Week 24 (for P030) showed no difference between gefapixant and placebo (Table 13 and Figure 8).

Although the gefapixant trials did not demonstrate a treatment effect on LCQ physical domain score, it should be noted that, if such an effect had been detected, the threshold to define a meaningful within-patient change is not established and would require supportive analysis, as discussed in Section 6.4.1.

Table 13. Change From Baseline in LCQ Physical Domain Score at Week 24, Trial P030 and Week 12, Trial P027 (FAS Population)

	Trial P030			Trial P027		
Statistic	**Placebo**	**Gefapixant 15 mg**	**Gefapixant 45 mg**	**Placebo**	**Gefapixant 15 mg**	**Gefapixant 45 mg**
N	406	404	399	216	223	212
Mean (SD) at baseline	3.9 (1.0)	3.9 (1.0)	3.9 (1.0)	3.8 (1.0)	3.9 (1.0)	3.9 (0.9)
Mean (SD) at Week 24 or 12*	4.8 (1.2)	4.9 (1.2)	5.0 (1.2)	4.5 (1.1)	4.7 (1.1)	4.7 (1.1)
Mean change from baseline (95% CI)**	0.93 (0.81, 1.04)	1.00 (0.88, 1.11)	1.06 (0.95, 1.18)	0.70 (0.56, 0.84)	0.74 (0.60, 0.87)	0.74 (0.60, 0.87)
Treatment difference (95% CI)***		0.07 (-0.08, 0.22)	0.13 (-0.02, 0.28)		0.04 (-0.14, 0.22)	0.04 (-0.15, 0.22)

Source: adqs.xpt; Table 4-29 in P030 CSR, Table 4-28 in CSR P027; Statistical Analyst.

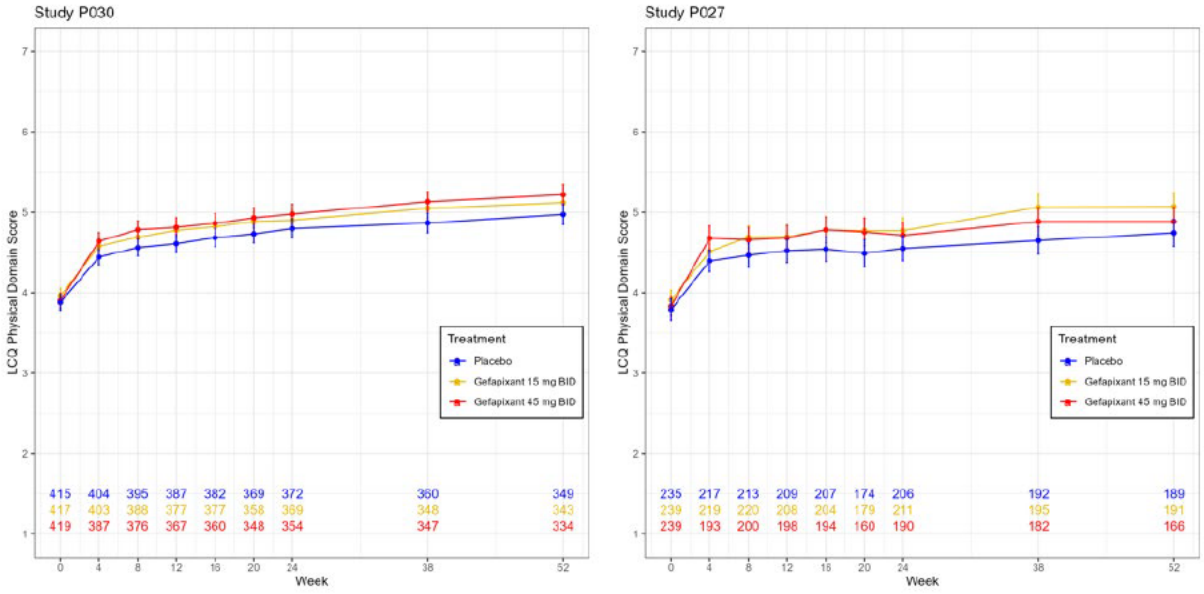
* Based on subjects with nonmissing values at baseline and Week 24 (P030) and Week 12 (P027).

** Based on the MMRM model.

*** The estimated difference is the treatment difference in model based mean change from baseline at Week 24 (P030) and Week 12 (P027).

Abbreviations: CI, confidence interval; CSR, clinical study report; FAS, full analysis set; LCQ, Leicester Cough Questionnaire; MMRM, mixed model repeated measures; N, number of subjects who had baseline and postbaseline assessments; SD, standard deviation

Figure 8. Mean (±95% CI) LCQ Physical Domain Score Over Time for Trials P030 and P027 (FAS Population)

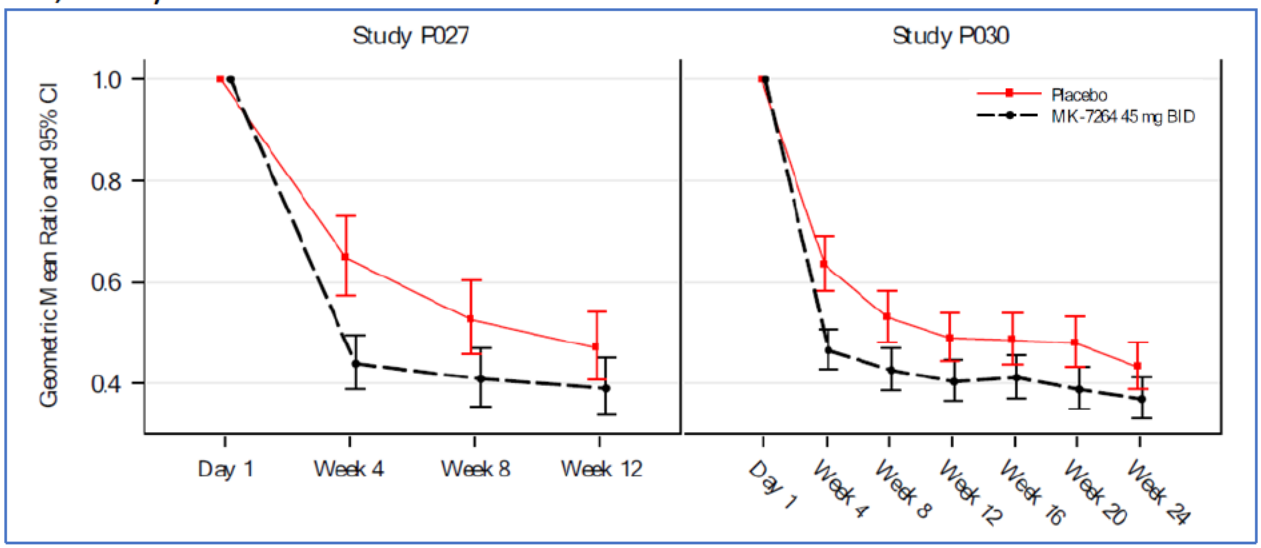


Source: Statistical analyst.
Abbreviations: BID, twice daily; FAS, full analysis set; LCQ, Leicester Cough Questionnaire; SE, standard error

3.1.3.7 Durability of Response

Results for 24-hour cough frequency over the main study period are shown for P030 and P027 in Figure 9. As cough frequencies were not recorded after the main study period ended (at Week 24 in P030 and Week 12 in P027), durability of treatment beyond these time points cannot be assessed. While there appeared to be treatment separation at Week 4, the placebo response continued to increase by Week 8 and tracked with the 45 mg arm at later weeks in both trials.

Figure 9. Twenty-Four-Hour Cough Frequency Over Time for Trials P027 and P030 (Full Analysis Set, Recount Data, MMRM)

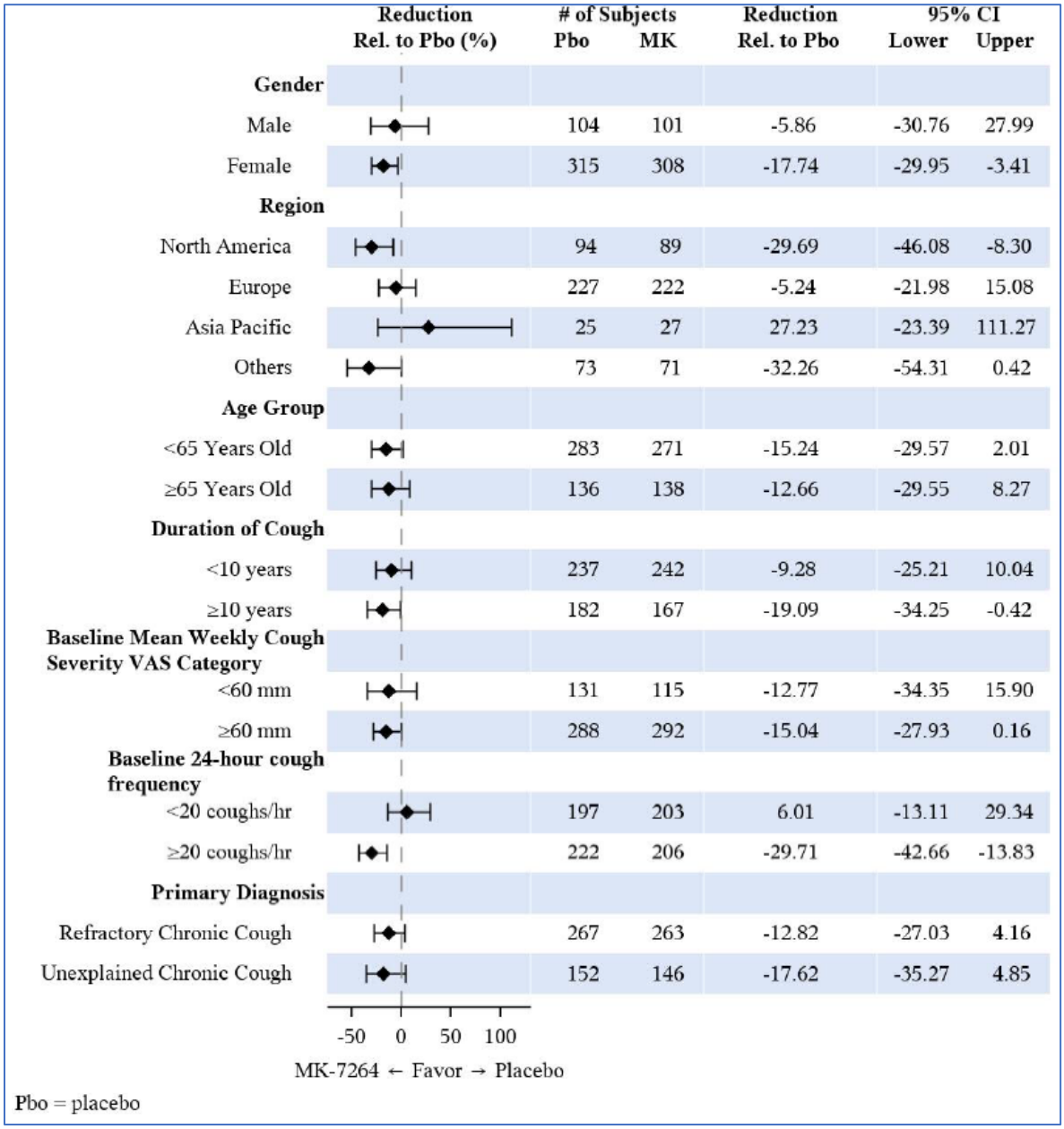


Source: Applicant's response to information request, 12 September 2023
Abbreviations: BID, twice daily; CI, confidence interval; MK-264, gefapixant

3.1.3.8 P030 and P027 Additional Analyses

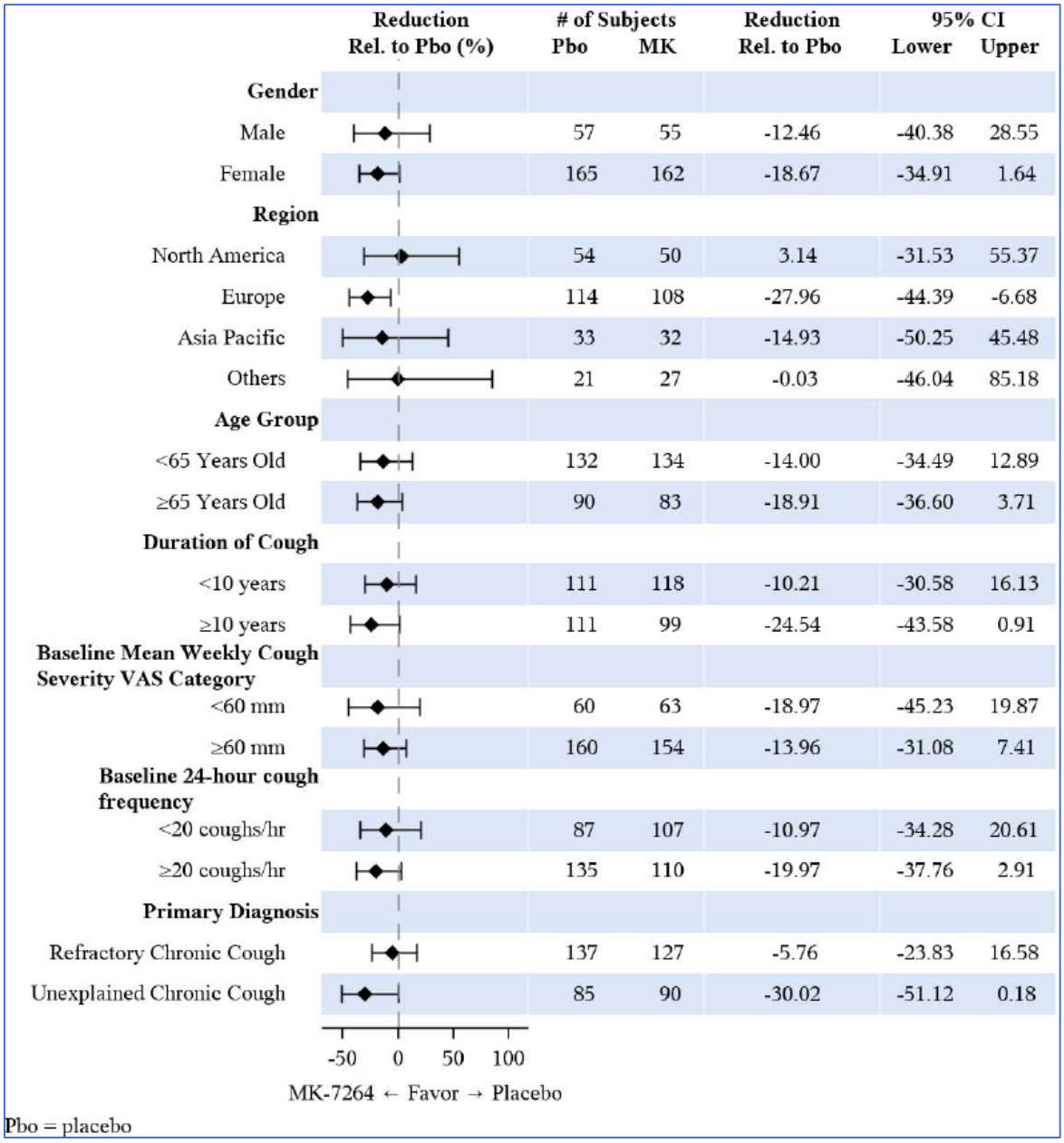
Forest plots showing subgroup analyses by demographic and baseline disease characteristics for the primary endpoint of 24-hour cough frequency at Week 24 for Trial P030 and Week 12 for Trial P027 are presented in Figure 10 and Figure 11, respectively. These results do not identify a subgroup of patients with increased responsiveness to gefapixant treatment.

Figure 10. Twenty-Four-Hour Cough Frequency by Demographic and Baseline Disease Characteristic Subgroups for Trial P030 at Week 24 (Full Analysis Set, Recount Data, MMRM)



Source: Applicant's response to information request, September 12, 2023, Figures 2-4 and 2-3.
Abbreviations: CI, confidence interval; MK and MK-7264, gefapixant 45 mg; VAS, visual analog scale

Figure 11. Twenty-Four-Hour Cough Frequency by Demographic and Baseline Disease Characteristic Subgroups for Trial P027 at Week 12 (Full Analysis Set, Recount Data, MMRM)



	Reduction Rel. to Pbo (%)	# of Subjects Pbo	# of Subjects MK	Reduction Rel. to Pbo	95% CI Lower	95% CI Upper
Gender						
Male		57	55	-12.46	-40.38	28.55
Female		165	162	-18.67	-34.91	1.64
Region						
North America		54	50	3.14	-31.53	55.37
Europe		114	108	-27.96	-44.39	-6.68
Asia Pacific		33	32	-14.93	-50.25	45.48
Others		21	27	-0.03	-46.04	85.18
Age Group						
<65 Years Old		132	134	-14.00	-34.49	12.89
≥65 Years Old		90	83	-18.91	-36.60	3.71
Duration of Cough						
<10 years		111	118	-10.21	-30.58	16.13
≥10 years		111	99	-24.54	-43.58	0.91
Baseline Mean Weekly Cough Severity VAS Category						
<60 mm		60	63	-18.97	-45.23	19.87
≥60 mm		160	154	-13.96	-31.08	7.41
Baseline 24-hour cough frequency						
<20 coughs/hr		87	107	-10.97	-34.28	20.61
≥20 coughs/hr		135	110	-19.97	-37.76	2.91
Primary Diagnosis						
Refractory Chronic Cough		137	127	-5.76	-23.83	16.58
Unexplained Chronic Cough		85	90	-30.02	-51.12	0.18

Pbo = placebo

Source: Applicant's response to Information Request, September 12, 2023, Figures 2-4 and 2-3.

Abbreviations: CI, confidence interval; MK and MK-7264, gefapixant 45 mg; VAS, visual analog scale

3.1.4 Efficacy Conclusions

In summary, the Applicant conducted two adequate and well-controlled pivotal trials, P030 and P027. Based on the prespecified MMRM analysis of the validated recount data, which the FDA considers the primary efficacy analysis, the primary endpoint of 24-hour cough frequency yielded a 15% to 17% relative reduction in geometric mean ratios with gefapixant 45 mg compared to placebo. The point estimate for the primary endpoint was similar in both trials but reached statistical significance in trial P030 only. Statistical significance does not, by itself, indicate whether the treatment effect corresponds

to a clinically meaningful benefit to patients. Because the primary endpoint results are challenging to interpret as presented (i.e., log transformation, geometric mean ratio, relative reduction), FDA conducted several post hoc analyses to facilitate our understanding of the reduction in cough frequency. Assessed a variety of ways, the reduction in cough frequency was consistently small. Further, the trial data provided no clear correlation between patients who reported improvement (per the PGIC) and reduction in cough frequency, no identifiable subgroup of patients who are more likely to respond to gefapixant, and no compelling additional support from the secondary endpoints related to cough (awake cough frequency, ≥30% reduction from baseline in 24-hour cough frequency). With a large placebo response and wide range of cough frequencies at baseline, we question if the small treatment difference in cough frequency is noticeable to patients. What constitutes a meaningful reduction in cough frequency is not established; therefore, we ask the Committee to consider whether the change in cough frequency with gefapixant will be perceptible and meaningful to patients.

Since the clinical meaningfulness of the primary endpoint results is not obvious, FDA has considered the supportive evidence from the secondary PRO endpoints to assess whether the small reduction in cough frequency is meaningful to patients. Based on the testing hierarchy in each trial, the LCQ total score is the only PRO endpoint that achieved statistical significance. The results from P030 showed a statistically significant, but numerically small, increase in the odds ratio for the proportion of patients who had a ≥1.3-point increase from baseline in the LCQ total score (range 3 to 21). While the odds ratio for the LCQ total score in P030 was statistically significant, the difference in the proportion of 'responders' was also small, and the absolute change from baseline in the LCQ total score compared to placebo was small (< 1 point) in relationship to the range of possible scores. Post hoc analyses of the LCQ physical domain, which has more relevance for informing regulatory decisions, showed no difference between gefapixant and placebo in the change from baseline in absolute score at Week 12 (P027) and Week 24 (P030).

Other PRO secondary endpoints (responder analyses on the CSD total score and Cough Severity VAS score) demonstrated small numerical increases in the odds ratios for a treatment response, but these were not controlled for multiplicity and did not have sufficient evidence to support the selected responder thresholds. Although the PRO results appear to favor gefapixant, the results must be interpreted with caution because 1) the measured absolute differences from placebo in the total score are small, and the difference in responders between treatment groups is small, 2) the degree of change in PRO scores corresponding to clinically meaningful improvements has not been established, 3) with the exception of the LCQ responder analysis in P030, none of these analyses are controlled for multiplicity; and 4) there is concern for potential unblinding due to taste disturbance (which occurred in up to 65% of treated subjects, discussed in Section 3.2.3). In light of these limitations and uncertainties, we ask the Committee to consider whether gefapixant's small treatment effect across PRO endpoints is meaningful to patients. Importantly, we ask the Committee to discuss whether the PRO data provide compelling evidence to inform the key question of whether the small reduction in cough frequency with gefapixant is clinically meaningful to patients.

3.2 Safety Issues

3.2.1 Sources of Data for Safety

To support the safety of gefapixant, the Applicant submitted safety data from the clinical efficacy trials (the pivotal trials P030 and P027, and the supportive phase 3b trials P042 and P043), as well as safety

results from a country-specific extended enrollment period of protocol P030, which occurred in China and is referred to as the China-specific P030 trial.

Because pivotal trials P030 and P027 were similar in design, enrolled the same population, had the same treatment arms, and had the same duration of treatment and safety follow-up, the assessment of safety was primarily based on pooled data from P030 and P027 through Week 52, which will be referred to as the Pooled Safety Dataset. All safety analyses were performed using the all-participants-as-treated (APaT) population, which includes all randomized subjects who received at least one dose of study intervention. The APaT population assigns treatment arm according to the treatment received rather than treatment assigned at randomization.

Supplemental analyses of safety results from the additional clinical trials with unique designs, populations, and durations (China-specific P030, P042, and P043) were conducted independently for each trial and compared to the primary safety analysis of the Pooled Safety Dataset. These analyses did not reveal safety signals that differed notably from the primary safety analysis results, and therefore this document will present the primary analysis of the Pooled Safety Dataset only.

3.2.2 Safety Summary

Overall Exposure

Across the clinical development program for gefapixant (phase 1, phase 2, and phase 3 trials), 3169 participants received at least 1 dose of active treatment. In the Pooled Safety Dataset from the pivotal trials, exposure to the gefapixant 45 mg dose was lower compared to the 15 mg dose and placebo. This was due to the high discontinuation rate due to adverse events (AEs) in the 45 mg arm, which mostly occurred in the first 8 weeks of treatment. Nonetheless, most subjects across treatment arms were exposed for at least 359 days, including 61.9% of subjects in the 45 mg treatment arm. Exposure data are listed in Table 14.

Table 14. Duration of Exposure, APaT Population, Pooled Safety Dataset

Variable	Gefapixant 15 mg N=686 n (%)	Gefapixant 45 mg N=683 n (%)	Placebo N=675 n (%)
Duration of exposure, days			
Mean (SD)	303.6 (116.8)	262.6 (144.2)	309.4 (113)
Median (Q1, Q3)	364 (355, 364.8)	363 (107.5, 364)	364 (360, 365)
Min, max	1, 418	1, 397	1, 396
Total exposure (person years)	570	491	572
Patients treated, by duration, n (%)			
<28 days	31 (4.5)	80 (11.7)	27 (4.0)
≥28 to <56 days	21 (3.1)	46 (6.7)	21 (3.1)
≥56 to <84 days	27 (3.9)	26 (3.8)	22 (3.3)
≥84 to <168 days	44 (6.4)	52 (7.6)	42 (6.2)
≥168 to <359 days	60 (8.7)	56 (8.2)	41 (6.1)
≥359 days	503 (73.3)	423 (61.9)	522 (77.3)

Source: FDA clinical data scientist.

Abbreviations: APaT, all participants as treated; N, number of subjects in treatment arm; n, number of subjects with given treatment duration; Q1, first quartile; Q3, third quartile; SD, standard deviation

Adequacy of the Safety Database

FDA and the International Conference on Harmonisation recommend that for products intended for long-term treatment of non-life-threatening conditions, 1500 subjects be exposed to the investigational product, with 300 to 600 exposed for 6 months and 100 exposed for 1 year; see the FDA guidance for industry: *Premarketing Risk Assessment* (2005). The safety database described above fulfills this recommendation, and no safety signals have been identified in the clinical development program to warrant a more extensive safety database than what was provided by the Applicant.

Deaths

Deaths were rare in the Pooled Safety Dataset, and no deaths occurred in subjects who received gefapixant 45 mg. The 2 deaths that occurred in subjects who received gefapixant 15 mg were unlikely to be related to study treatment. There is no signal for increased mortality with gefapixant.

Serious Adverse Events

Serious adverse events (SAEs) occurred in approximately 6% of subjects in the Pooled Safety Dataset and were numerically similar across treatment arms. Review of the System Organ Classes (SOC) and Preferred Terms (PTs) for the SAEs did not reveal a pattern suggestive of a relationship between the SAEs and gefapixant treatment.

Treatment Discontinuations Due to Adverse Effects

There was a dose-dependent increase in AEs leading to treatment discontinuation when comparing placebo to gefapixant. In the gefapixant 45 mg arm, 22.3% of subjects discontinued treatment compared to 7.9% in the 15 mg arm and 5.6% in the placebo arm. The imbalance in discontinuations with gefapixant was primarily due to AEs in the Gastrointestinal Disorders and Nervous System Disorders SOCs, which reflects the common occurrence of taste-related AEs and oropharyngeal AEs. Taste-related AEs led to early treatment discontinuation in 14% of subjects in the 45 mg dose group compared to 0.3% and 1.3% of subjects in the placebo and 15 mg dose groups, respectively. Taste-related AEs will be discussed separately below.

Severe Adverse Events

In the pivotal trials, severe AEs were defined as AEs that result in an inability to perform daily activities. Severe dysgeusia events and severe cough events occurred more frequently with gefapixant 45 mg compared to placebo; however, the rates of these severe AEs were low with 1.6% and 0.7% of subjects in the gefapixant 45 mg arm experiencing severe dysgeusia and severe cough, respectively. Cough may occur more frequently in the gefapixant 45 mg arm because of potentially related upper respiratory/oropharyngeal AEs, including dry mouth/throat and throat irritation, exacerbating the underlying chronic cough.

Common Adverse Events

Treatment-emergent adverse events (TEAE) were common and occurred in more subjects who received gefapixant 45 mg (88.9%) compared to both 15 mg (81.5%) and placebo (79.0%). This difference was largely driven by taste-related AEs. Taste-related AEs will be discussed in more detail below. Other common (≥5%) TEAEs occurring more with gefapixant than placebo can be generally characterized as

impacting the upper airway (cough, oropharyngeal pain, upper respiratory infection) or gastrointestinal (GI) tract (dry mouth, nausea, diarrhea). The prevailing pattern of TEAEs impacting the upper airway and upper GI tract is likely related to gefapixant's mechanism of action, which targets sensory nerves in these anatomic locations. While the incidence of upper airway and upper GI tract AEs is dose-dependent, they can be monitored in a clinical setting.

Vital Signs and Laboratory Findings

No clinically significant changes in vital signs or clinical laboratory findings were identified.

Safety Analyses by Demographic Subgroups

Safety analyses were conducted for demographic subgroups, including by age, gender, race, and geographic region. No notable differences impacting the risk-benefit determination in a specific subgroup were observed.

Drug-Specific Safety Issues

Certain safety issues underwent targeted analysis given the mechanism of action and known safety profile of gefapixant. These issues included taste-related AEs, renal and urinary events, lower respiratory tract infections and pneumonias, hypersensitivity, obstructive sleep apnea, and hepatotoxicity. Upon review of these issues, no specific safety signal was observed except for taste-related AEs, discussed further below.

3.2.3 Safety Issues in Detail

Taste-Related AEs

Changes in taste sensation are a known pharmacologic effect of P2X3 inhibitors. In the gefapixant pivotal trials, taste-related AEs, including dysgeusia, ageusia, taste disorder, hypogeusia, and hypergeusia, were predefined at the study design stage as safety endpoints of special interest. There was a clear dose-dependent increase in taste-related AEs, which occurred in 65.4% of subjects in the gefapixant 45 mg group, with a large majority of cases being mild or moderate in intensity. These data are shown in Table 15.

Table 15. Taste-Related AEs Occurring in ≥1 Subject in Any Treatment Arm, ApAT, Pooled Safety Dataset

Preferred Term Maximum Intensity	Placebo N=675 n (%)	Gefapixant 15 mg N=686 n (%)	Gefapixant 45 mg N=683 n (%)
At least 1 taste-related AE	47 (7.0)	120 (17.5)	447 (65.4)
Mild	41 (6.1)	93 (13.6)	289 (42.3)
Moderate	6 (0.9)	25 (3.6)	141 (20.6)
Severe	0	2 (0.3)	17 (2.5)
Ageusia	6 (0.9)	16 (2.3)	100 (14.6)
Dysgeusia	36 (5.3)	78 (11.4)	281 (41.1)
Hypergeusia	2 (0.3)	2 (0.3)	5 (0.7)
Hypogeusia	4 (0.6)	22 (3.2)	73 (10.7)
Taste disorder	3 (0.4)	10 (1.5)	61 (8.9)

Source: Integrated Summary of Safety Table 5.3.5.3.3-cough: 31; confirmed by clinical reviewer.
Abbreviations: AE, adverse event; APaT, all participants as treated; N, number of subjects in treatment arm; n, number of subjects with adverse event

The taste-related AEs associated with gefapixant 45 mg are characterized by rapid onset (median 2 days) and mean duration of 204 days (median 194 days). Resolution of taste-related AEs occurs in 96% of subjects in the gefapixant 45 mg arm, primarily after treatment discontinuation, although ~25% of events resolved on or before the last dose. In the subjects for whom resolution was not documented (i.e., taste-related AE had a missing AE end date), it is unclear whether the taste-related AE was truly ongoing or whether the AE end date was missing for other reasons (e.g., failure to document end date). Given gefapixant's mechanism of action and relatively short half-life, it is unlikely that the missing AE end date in this small number of cases is indicative of a prolonged or irreversible effect.

To assess for potential systemic sequelae of taste-related AEs, events including loss of appetite, weight loss, and dehydration were evaluated. Decreased appetite occurred most frequently in the gefapixant 45 mg arm (3.7%) compared to 15 mg (0.6%) and placebo (1.0%). However, there was no clear association between gefapixant and more substantial systemic AEs such as weight loss and dehydration.

Overall, gefapixant causes frequent taste-related AEs in a dose-dependent fashion. Decreased appetite also occurs at lower frequency. These taste-related AEs frequently led to treatment discontinuation (16.1% and 11.9% of subjects receiving 45 mg in P030 and P027, respectively), most often occurring within the first 4-8 weeks of therapy. However, these AEs appear to be readily reversible with or without cessation of therapy and do not appear to result in more concerning systemic effects like weight loss or dehydration.

4 Benefit-Risk Framework

Benefit-Risk Framework

Disclaimer: This pre-decisional Benefit-Risk Framework does not represent the FDA's final benefit-risk assessment or regulatory decision.

	Evidence and Uncertainties	Comments to the Advisory Committee
Analysis of Condition	CC is a common condition, with a prevalence of 5-10% of adults. CC is not a life-threatening condition, but it can have negative, long-term impacts on quality of life. The natural history of CC is poorly understood; in some patients, it presents as daily cough lasting years or decades, and in others, the course is relapsing and remitting.	CC is a common condition that impacts quality of life but not mortality. As a chronic condition, long-term treatment is anticipated.
Current Treatment Options	There are no currently approved therapies for CC. Off-label treatments include opioids, neuroleptics, and local anesthetics; these treatment options carry risks and have limited evidence supporting treatment benefit.	There are no currently approved treatments for CC, indicating that this is an area of unmet medical need.
Benefits	• The results of 2 adequate and well-controlled trials showed a relative reduction in the geometric mean ratio reflecting change from baseline in cough frequency of 15 to 17% with gefapixant 45 mg treatment compared to placebo. • Post hoc analysis of the median absolute change from baseline in 24-hour cough frequency showed a decrease of 1 to 2 cough frequency with gefapixant 45 mg compared to placebo. • The clinical meaningfulness of the observed small relative reduction in cough frequency is uncertain. • Secondary endpoints of awake cough frequency and responders of 30% reduction in cough frequency did not provide additional insight to the clinical meaningfulness of the cough reduction with gefapixant. • There was a significant difference in responders (≥1.3-point increase from baseline in LCQ total score at Week 24) with gefapixant compared to placebo in P030, but the difference in the proportion of responders in LCQ total score with gefapixant compared to placebo is small. • Although not controlled for multiplicity, analyses for the CSD and Cough Severity VAS showed numerical increases in responders with gefapixant compared to placebo, and there is not sufficient evidence to support the selected responder thresholds. The differences between gefapixant and placebo in proportion of responders in the CSD and Cough Severity VAS are small.	• Results from the clinical program show a small reduction in cough frequency with gefapixant relative to placebo and a small treatment effect across PROs. It is unclear whether the small reduction in cough frequency with gefapixant and the results from the PROs demonstrate a clinically meaningful benefit to patients. • Secondary endpoints related to cough frequency do not appear to provide additional support to the clinical meaning of the primary endpoint results. • The observed PRO results have limitations impacting their ability to inform the clinical meaningfulness of the reduction in cough frequency. • Potential unblinding due to taste disturbance may have affected the cough and PRO results. • The clinical meaningfulness of the benefits of gefapixant for treatment of CC are unclear.

	Evidence and Uncertainties	Comments to the Advisory Committee
	• There is a large placebo response for all the efficacy endpoints in the clinical program. • An additional uncertainty regarding efficacy is the potential occurrence of unblinding due to the high incidence of taste disturbance.	
Risks and Risk Management	• Taste disturbances occurred in 65% of patients who received gefapixant and led to early treatment discontinuation in 14%. • Taste disturbance was generally reversible (in 96% of patients), primarily upon treatment discontinuation, although ~25% resolved on treatment.	Taste disturbance is the primary risk with gefapixant treatment impacting tolerability but can be mitigated with labeling and treatment discontinuation.

Abbreviations: CC, chronic cough; LCQ, Leicester Cough Questionnaire; PRO, patient-reported outcome

Summary of Benefit-Risk

For a drug to be approved for marketing in the United States, FDA must determine that the drug is effective and that its expected benefits outweigh its potential risks to patients. A benefit-risk assessment for gefapixant requires careful consideration of the evidence and remaining uncertainties about its key benefits (as demonstrated in the development program) and potential key risks, as well as the ability to mitigate such risks. This assessment should consider the unmet need for patients with CC, while bearing in mind that the condition is not life-threatening or rare and may require chronic long-term therapy.

The key issue for the benefit-risk assessment is uncertainty about the clinical meaningfulness about the benefit of gefapixant, given the small reduction in cough frequency in the setting of a high placebo response and uncertainties surrounding the interpretation of the limited supporting evidence from PROs, which also showed a small treatment effect. The safety profile is notable for the high frequency of taste disturbance, but these AEs are readily reversible. While the gefapixant clinical program did not identify serious safety concerns from its use, a favorable benefit-risk profile cannot be concluded if there is no meaningful benefit of the therapy to patients.

Point(s) to Consider: Given the small reduction in cough frequency with gefapixant treatment, the observed results from patient-reported outcomes, and the potential unblinding of patients due to taste disturbance, do the available data support that treatment with gefapixant provides a clinically meaningful benefit for adult patients with refractory or unexplained chronic cough?

5 References

Literature and Websites

Birring, SS, B Prudon, AJ Carr, SJ Singh, MD Morgan, and ID Pavord, 2003, Development of a symptom specific health status measure for patients with chronic cough: Leicester Cough Questionnaire (LCQ), Thorax, 58(4):339-343.

Canning, BJ, AB Chang, DC Bolser, JA Smith, SB Mazzone, L McGarvey, and CEC Panel, 2014, Anatomy and neurophysiology of cough: CHEST Guideline and Expert Panel report, Chest, 146(6):1633-1648.

FDA-NIH, 2016, BEST (Biomarkers, EndpointS, and other Tools) Resource, accessed, .

Ford, AP, 2012, In pursuit of P2X3 antagonists: novel therapeutics for chronic pain and afferent sensitization, Purinergic Signal, 8(Suppl 1):3-26.

Gibson, P, G Wang, L McGarvey, AE Vertigan, KW Altman, SS Birring, and CEC Panel, 2016, Treatment of Unexplained Chronic Cough: CHEST Guideline and Expert Panel Report, Chest, 149(1):27-44.

Morice, AH, AD Jakes, S Faruqi, SS Birring, L McGarvey, B Canning, JA Smith, SM Parker, KF Chung, K Lai, ID Pavord, J van den Berg, WJ Song, E Millqvist, MJ Farrell, SB Mazzone, P Dicpinigaitis, and R Chronic Cough, 2014a, A worldwide survey of chronic cough: a manifestation of enhanced somatosensory response, Eur Respir J, 44(5):1149-1155.

Morice, AH, E Millqvist, MG Belvisi, K Bieksiene, SS Birring, KF Chung, RW Dal Negro, P Dicpinigaitis, A Kantar, LP McGarvey, A Pacheco, R Sakalauskas, and JA Smith, 2014b, Expert opinion on the cough hypersensitivity syndrome in respiratory medicine, Eur Respir J, 44(5):1132-1148.

Morice, AH, E Millqvist, K Bieksiene, SS Birring, P Dicpinigaitis, C Domingo Ribas, M Hilton Boon, A Kantar, K Lai, L McGarvey, D Rigau, I Satia, J Smith, WJ Song, T Tonia, JWK van den Berg, MJG van Manen, and A Zacharasiewicz, 2020, ERS guidelines on the diagnosis and treatment of chronic cough in adults and children, Eur Respir J, 55(1).

Nguyen, AM, J Schelfhout, D Muccino, ED Bacci, C La Rosa, M Vernon, and SS Birring, 2022, Leicester Cough Questionnaire validation and clinically important thresholds for change in refractory or unexplained chronic cough, Ther Adv Respir Dis, 16:17534666221099737.

Raj, AA, DI Pavord, and SS Birring, 2009, Clinical cough IV: what is the minimal important difference for the Leicester Cough Questionnaire?, Handb Exp Pharmacol, (187):311-320.

Guidances for Industry

FDA, 2022, Guidance for Industry; Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments.

6 Appendix

6.1 Pivotal Trials P030 and P027: Additional Information

6.1.1 Restricted Therapies During the Course of the Trials

- Opioids for treatment of cough (use for other indications was permitted if treatment regimen was stable)
- Pregabalin, gabapentin, amitriptyline, or nortriptyline for treatment of cough (use for other indications was permitted if treatment regimen was stable)
- Dextromethorphan, guaifenesin, benzonatate, and any other over-the-counter or prescription medication for the treatment of cough
- Treatments for conditions associated with chronic cough, such as GERD, asthma, UACS, or nonasthmatic eosinophilic bronchitis, are permitted for subjects who had been on a stable treatment regimen for at least 2 months
- Nonpharmacologic treatments (e.g., physiotherapy, speech and language therapy)
- Angiotensin converting enzyme inhibitors

6.1.2 Subject Disposition

Table 16 and Table 17 show subject disposition through Week 52 in Trials P030 and P027, respectively.

Table 16. P030 Subject Disposition Through Week 52, All Subjects as Randomized*

Disposition Outcome	Gefapixant 15 mg N=442 n (%)	Gefapixant 45 mg N=439 n (%)	Placebo N=436 n (%)
Discontinued study	74 (16.7)	84 (19.1)	54 (12.4)
Death	1 (0.2)	0	0
Lost to follow-up	2 (0.5)	5 (1.1)	6 (1.4)
Other	1 (0.2)	2 (0.5)	0
Physician decision	0	3 (0.7)	1 (0.2)
Screen failure	2 (0.5)	0	1 (0.2)
Withdrawal by subject	68 (15.4)	74 (16.9)	46 (10.6)
Discontinued treatment	113 (25.6)	156 (35.5)	85 (19.5)
Adverse event	39 (8.8)	100 (22.8)	25 (5.7)
Death	1 (0.2)	0	0
Lost to follow-up	2 (0.5)	4 (0.9)	5 (1.1)
Noncompliance with treatment	1 (0.2)	2 (0.5)	3 (0.7)
Other	2 (0.5)	0	2 (0.5)
Physician decision	2 (0.5)	5 (1.1)	3 (0.7)
Pregnancy	1 (0.2)	0	0
Withdrawal by subject	65 (14.7)	45 (10.3)	47 (10.8)
Did not receive any treatment	2 (0.5)	0	1 (0.2)

Source: FDA Clinical Data Scientist.

* Minor differences from the Applicant's disposition analysis in the percentage of subjects who discontinued study drug are due to inclusion of subjects who did not receive any study drug in the calculations for this table.

Abbreviations: N, number of subjects in treatment arm; n, number of subjects in specified population or group

Table 17. P027 Subject Disposition Through Week 52, All Subjects as Randomized*

Disposition Outcome	Gefapixant 15 mg N=244 n (%)	Gefapixant 45 mg N=244 n (%)	Placebo N=244 n (%)
Discontinued study	44 (18)	60 (24.6)	45 (18.4)
Death	1 (0.4)	0	2 (0.8)
Lost to follow-up	1 (0.4)	1 (0.4)	2 (0.8)
Other	0	0	1 (0.4)
Physician decision	3 (1.2)	3 (1.2)	2 (0.8)
Screen failure	0	1 (0.4)	1 (0.4)
Withdrawal by subject	39 (16)	55 (22.5)	37 (15.2)
Discontinued treatment	57 (23.4)	96 (39.3)	60 (24.6)
Adverse event	15 (6.1)	52 (21.3)	13 (5.3)
Death	1 (0.4)	0	2 (0.8)
Lost to follow-up	1 (0.4)	1 (0.4)	2 (0.8)
Noncompliance with treatment	1 (0.4)	2 (0.8)	0
Other	1 (0.4)	0	1 (0.4)
Physician decision	0	0	1 (0.4)
Pregnancy	0	1 (0.4)	0
Withdrawal by subject	38 (15.6)	40 (16.4)	41 (16.8)
Did not receive any treatment	0	1 (0.4)	1 (0.4)

Source: FDA Clinical Data Scientist.

*Minor differences from the Applicant's disposition analysis in the percentage of subjects who discontinued study drug are due to inclusion of subjects who did not receive any study drug in the calculations for this table.

Abbreviations: N, number of subjects in treatment arm; n, number of subjects in specified population or group

6.1.3 Summary of Primary Endpoint Sensitivity and Post Hoc Analyses

Several sensitivity analyses were conducted on the primary analysis, which are summarized in Table 18 and described below. The primary analysis results are included in the first row for comparison. For the Applicant's proposed dose, gefapixant 45 mg, compared to placebo, all sensitivity analyses had similar point estimates to the primary analysis. There were two differences in statistical significance in sensitivity and post hoc analyses compared to primary analyses: for the MI+ANCOVA analysis for P027, this post hoc result was statistically significant; for the J2R-multiple imputation analysis for P030, this sensitivity analysis was not statistically significant. Statistical significance for all other sensitivity analyses were consistent with the primary analysis results.

Table 18. Sensitivity Analyses Results for Estimated Relative Reduction (%) in 24-Hour Cough Frequency at Week 24 for Trial P030 (Week 12 for Trial P027)

	P030		P027	
Sensitivity Analysis	**Gefapixant 15 mg vs. Placebo**	**Gefapixant 45 mg vs. Placebo**	**Gefapixant 15 mg vs. Placebo**	**Gefapixant 45 mg vs. Placebo**
Primary analysis, MMRM	-2.7% (-15.6, 12.1)	-14.6%[1] (-26.0, -1.5)	-3.1% (-14.5, 24.4)	-17.0% (-31.5, 0.6)
J2R – Multiple imputation	-2.6 (-15.4, 12.3)	-12.6 (-24.4, 0.9)	2.4 (-13.8, 21.6)	-14.9 (-30.6, 4.4)
J2R – Pattern mixture	-2.4 (-13.7, 10.5)	-12.6 (-22.5, -1.3)	2.9 (-13.5, 22.4)	-15.3 (-28.6, 0.5)
MI+ANCOVA[2]	-0.9% (-14.1, 14.3)	-13.1% (-24.5, -0.1)	-1.2% (-15.8, 21.6)	-17.1% (-31.2, -0.1)
MI+ANCOVA in subjects with baseline and at least one postbaseline value[3]	-2.3% (-15.2, 12.6)	-14.5% (-25.8, -1.6)	2.1% (-15.2, 23.9)	-17.1% (-31.7, 0.5)
Primary analysis with 2 placebo outliers at baseline removed[4]	n/a	n/a	4.2 (-13.6, 25.8)	-16.2 (-30.8, 1.5)

Source: Statistical analyst. J2R Source: 13 October 2023 Information Response.

[1] Tipping point analysis supported robustness of this result.

[2] Results for Applicant's MI + ANCOVA (not prespecified) are described in the Appendix.

[3] Results for modification of Applicant's MI + ANCOVA in subjects with a baseline value and at least one postbaseline value (not prespecified), described in the Appendix.

[4] Subject who was removed had an ID number of (b) (6).

Abbreviations: ANCOVA, analysis of covariance; MI, multiple imputation; n/a, not applicable; J2R, jump to reference

6.1.4 Jump to Reference Sensitivity Analyses for Primary Results

Jump to reference (J2R) missing data sensitivity analyses, which test for the missing at random assumption were conducted for the primary analysis and are summarized in Table 18 (J2R - multiple imputation method was prespecified in the original statistical analysis plan (SAP); J2R – pattern mixture model approximation was requested by the FDA). The only sensitivity analysis that was not consistent with the primary analysis was the J2R-multiple imputation analysis for P030, where the sensitivity analysis was not statistically significant.

6.1.5 Tipping Point Analysis and Robustness of Primary Results, P030

For the significant primary efficacy result in Study P030 at the 45 mg dose, the Applicant conducted a tipping point analysis as defined in their SAP for the original submission and using the resubmitted, validated data. The same amount of percentage worsening was applied to both the gefapixant 15 mg BID group and the gefapixant 45 mg BID group (which could be different from the percentage worsening applied to the placebo group) and the analysis model was run on all three groups of subjects, but only the comparison between gefapixant 45 mg BID and placebo was presented. This approach was consistent with the prespecified tipping point analysis described in the Applicant's original SAP.

The point where the primary comparison of the gefapixant group versus the placebo group was no longer statistically significant appeared between worsening percentage of 7% and 8% in the gefapixant group. Since the penalty at the tipping point generates lower cough reduction for gefapixant-treated subjects with missing data than for placebo-treated subjects with missing data, the penalty might be considered implausibly high, and the tipping point analysis is interpreted as supporting robustness of the data.

Table 19. Twenty-Four-Hour Cough Frequency at Week 24 Tipping Point Analysis Using Multiple Imputation, Trial P030

Worsening (%)† Applied to the Imputed Data in Placebo Group	Worsening (%)† Applied to the Imputed Data in Gefapixant 45 mg BID Group									
	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
3%	0.0286	0.0307	0.0329	0.0351	0.0375	0.0400	0.0426	0.0454	0.0482	0.0512*
2%	0.0301	0.0323	0.0345	0.0369	0.0394	0.0420	0.0447	0.0476	0.0505*	0.0536*
1%	0.0317	0.0339	0.0363	0.0388	0.0414	0.0441	0.0469	0.0499	0.0530*	0.0562*
0%	0.0333	0.0357	0.0381	0.0407	0.0434	0.0463	0.0492	0.0523*	0.0555*	0.0588*
-1%	0.0351	0.0375	0.0401	0.0428	0.0456	0.0486	0.0516*	0.0548*	0.0582*	0.0617*
-2%	0.0369	0.0395	0.0422	0.0450	0.0479	0.0510*	0.0542*	0.0575*	0.0610*	0.0646*

Source: Applicant, Information Request, August 31, 2023

* p>0.05.

† x% worsening indicates the value is multiplied by (100+x)%. For the placebo cases with x<0, this means an improvement. P-Values based on the estimated treatment difference between the gefapixant 45 mg BID and placebo groups.

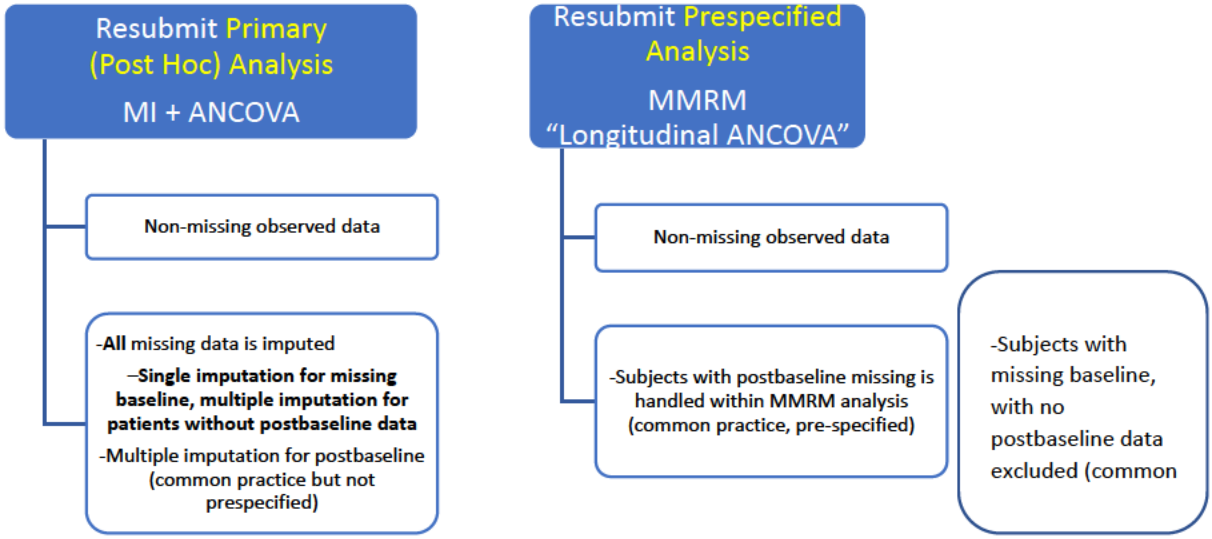
Abbreviation: BID, twice daily

6.1.6 Applicant's Proposed Efficacy Analysis (Not Prespecified)

For this resubmission, the Applicant proposed the use of a new analysis method as their primary analysis. This post hoc "MI + ANCOVA" approach imputed missing baseline values and/or missing postbaseline values. Specifically, there was a single imputation for missing baseline values based on gender and region; multiple imputations for the postbaseline (m=50 imputed datasets) were conducted for all follow-up visits using treatment, gender, region, and the other follow-up visits as covariates. Once these imputations were implemented, the ANCOVA model was conducted on the resultant data. This MI + ANCOVA approach used the entire FAS population by imputing all missing values. Following imputation, an ANCOVA model was conducted at Week 24, adjusting for covariates of treatment, baseline, region, and gender.

To further explore how the single imputation for patients without a baseline value influenced this analysis, we added an analysis similar to "MI + ANCOVA" as a sensitivity analysis, where multiple imputations for the postbaseline (m=50 imputed datasets) were conducted for all follow-up visits using treatment, gender, region, and the other follow-up visits as covariates, but patients with no baseline value and patients who had a baseline value but no postbaseline values were excluded. The key features of the Applicant's post hoc analysis and the prespecified primary analysis are characterized in Figure 12.

Figure 12. Primary Analyses for Primary Endpoint, 24-Hour Cough Frequency, P030 and P027 (FAS Population)



Source: Statistical reviewer.

Abbreviations: ANCOVA, analysis of covariance; FAS, full analysis set; MMRM, mixed model repeated measures

For P030, there were a total of 16 additional subjects with missing baseline (2 in placebo, 9 in 15 mg, and 5 in 45 mg gefapixant arms) and a total of 54 additional subjects with no postbaseline data (13 in placebo, 16 in 15 mg, and 25 in 45 mg gefapixant arms). This yielded a slightly lower point estimate of 13.1% reduction (95% CI, -24.5 to -0.1) with a p-value of 0.048 at the 45 mg dose (Table 20). Results at the 15 mg dose were not significant.

For P027, of the 63 additional subjects included, there were 25 additional subjects with missing baseline (11 in placebo, 9 in 15 mg, and 5 in 45 mg gefapixant arms) and 38 additional subjects with no postbaseline data (10 in placebo, 8 in 15 mg, and 20 in 45 mg gefapixant arms). This analysis yielded a slightly higher point estimate of 17.1% reduction (95% CI, -31.2 to -0.1) with a p-value of 0.049 at the 45 mg dose, which was not significant in the prespecified analysis (Table 21). Results at the 15 mg dose were not significant.

Table 20. Analysis of 24-Hour Cough Frequency at Week 24, Baseline and Postbaseline Imputation + ANCOVA, Trial P030 (Full Analysis Set; Recount Data)

Treatment	N	Estimated Relative Reduction (%) and (95% CI)*	p-Value
Placebo	435		
Gefapixant 15 mg BID	440	-0.9 (-14.1, 14.3)	0.900
Gefapixant 45 mg BID	439	-13.1 (-24.5, -0.1)	0.048

Source: adeff.xpt, adsl.xpt (original), adbase.xpt; Table 4-1 in study report validated by the statistical analyst.

*The estimated relative reduction (relative to placebo) is calculated by 100 (e^{DIFF} -1). DIFF is the treatment difference in change from baseline at Week 24 based on the log-transformed data.

Abbreviations: ANCOVA, analysis of covariance; BID, twice daily; CI, confidence interval; N, number of subjects included in the analysis

Table 21. Analysis of 24-Hour Cough Frequency at Week 12, Baseline and Postbaseline Imputation + ANCOVA, Trial P027 (Full Analysis Set; Recount Data)

Treatment	N	Estimated Relative Reduction (%) and (95% CI)*	p-Value
Placebo	243		
Gefapixant 15 mg BID	244	1.2 (-15.8, 21.6)	0.899
Gefapixant 45 mg BID	243	-17.1 (-31.2, -0.1)	0.049

Source: adeff.xpt, adsl.xpt (original), adbase.xpt; Table 4-1 in study report validated by the statistical analyst.
*The estimated relative reduction (relative to placebo) is calculated by 100 (e^{DIFF} -1). DIFF is the treatment difference in change from baseline at Week 12 based on the log-transformed data.
Abbreviations: ANCOVA, analysis of covariance; BID, twice daily; CI, confidence interval; N, number of subjects included in the analysis

An additional analysis similar to the Applicant's (not prespecified) primary analysis, "MI+ANCOVA" but excluding patients without baseline data and patients with baseline but not postbaseline data was conducted to assess how much the exclusion of subjects with missing baseline or with no postbaseline data contributed to the MI+ANCOVA results compared to use of multiple imputation on the remaining patients (which is the same population as the prespecified primary analysis). Results for this modified analysis in P030 were a relative reduction of 14.5% (95% CI, -25.8, -1.6) and p-value of 0.029 at the 45 mg dose (Table 22). For P027, this analysis yielded a slightly higher point estimate of 17.1% reduction (95% CI, -31.2 to 0.1) with a p-value of 0.057 at the 45 mg dose (Table 23). These results are very similar to the prespecified primary analysis, indicating that the reason for a different p-value result in the MI+ANCOVA analysis in P027 for the 45 mg dose arm compared to placebo was likely due to the additional patients included, and not due to use of the multiple imputation technique for addressing missing data. Results at the 15 mg dose were not significant.

Table 22. Analysis of 24-Hour Cough Frequency at Week 24, Subjects With a Baseline Value and at Least One Postbaseline Value, MI + ANCOVA, Trial P030 (Full Analysis Set; Recount Data)

Treatment	N	Estimated Relative Reduction (%) and (95% CI)*	p-Value
Placebo	419		
MK-7264 15 mg BID	415	-2.3 (-15.2, 12.6)	0.750
MK-7264 45 mg BID	409	-14.5 (-25.8, -1.6)	0.029

Source: adeff.xpt, adsl.xpt (original), adbase.xpt.
*The estimated relative reduction (relative to placebo) is calculated by 100 (e^{DIFF} -1). DIFF is the treatment difference in change from baseline at Week 12 based on the log-transformed data.
Abbreviations: ANCOVA, analysis of covariance; BID, twice daily; CI, confidence interval; N, number of subjects included in the analysis

Table 23. Analysis of 24-Hour Cough Frequency at Week 12, Subjects With a Baseline Value and at Least One Postbaseline Value, MI + ANCOVA, Trial P027 (Full Analysis Set; Recount Data)

Treatment	N	Estimated Relative Reduction (%) and (95% CI)*	p-Value
Placebo	222		
MK-7264 15 mg BID	227	2.5 (-15.2, 23.9)	0.802
MK-7264 45 mg BID	217	-17.1 (-31.7, 0.5)	0.057

Source: adeff.xpt, adsl.xpt (original), adbase.xpt.
* The estimated relative reduction (relative to placebo) is calculated by 100 (e^{DIFF} -1). DIFF is the treatment difference in change from baseline at Week 12 based on the log-transformed data.
Abbreviations: ANCOVA, analysis of covariance; BID, twice daily; CI, confidence interval; N, number of subjects included in the analysis

The final sensitivity analysis assessed 24-hour cough frequency at Week 12 using the prespecified primary analysis with the one extreme outlier removed (Table 24).

Table 24. Analysis of 24-Hour Cough Frequency, Trial P027 With Two Placebo Outliers Excluded (Full Analysis Set, Recount Data, MMRM)

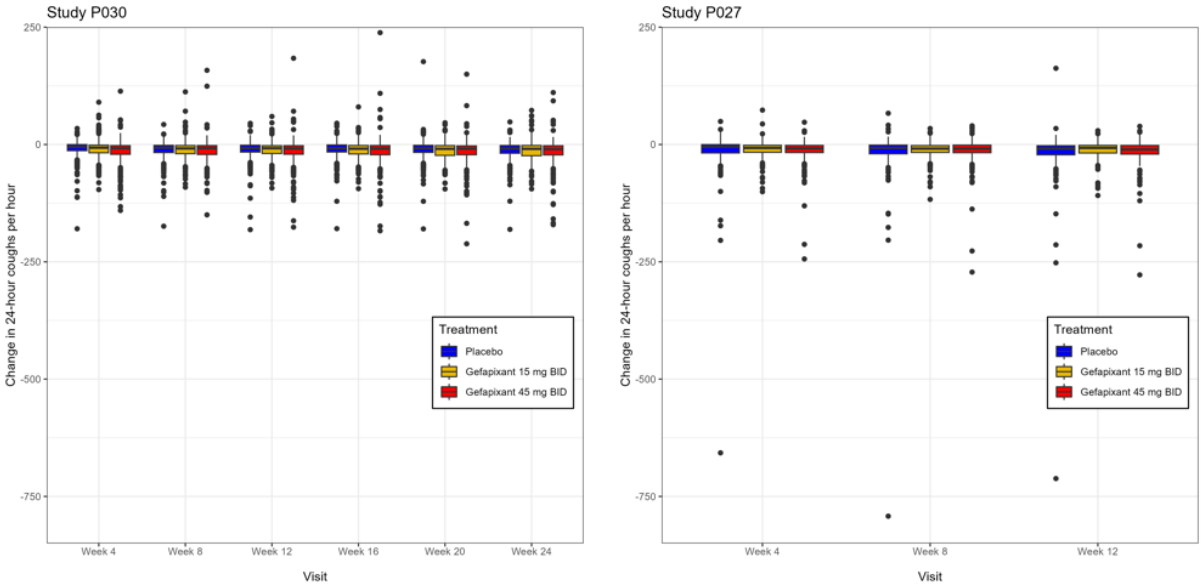
Statistic	Placebo	Gefapixant 15 mg	Gefapixant 45 mg
N	220	227	217
Geometric mean at baseline*	22.8	20.9	18.9
Geometric mean at Week 12*	10.3	10.2	7.4
Geometric mean ratio*	0.45	0.49	0.39
Model based geometric mean ratio (95% CI)**	0.47 (0.40, 0.54)	0.49 (0.42, 0.56)	0.39 (0.34, 0.45)
Relative reduction (%) in geometric mean ratio (95% CI)***		4.23 (-13.62, 25.77)	-16.19 (-30.80, 1.52)
p-value		0.665	0.071

Source: adeff.xpt; statistical analyst.
Subject IDs of those excluded: (b) (6).
* Based on subjects with nonmissing values at baseline and Week 12.
** Based on the MMRM model with change from baseline in log-transformed 24-hour cough frequency at each postbaseline visit (up to Week 12) as the response.
*** The estimated relative reduction (relative to placebo) is calculated by 100 (e^{DIFF} -1). DIFF is the treatment difference in change from baseline at Week 12 based on the log-transformed data.
Abbreviations: CI, confidence interval; N, number of subjects included in the analysis; MMRM, mixed model repeated measures

6.1.7 Exploratory Analysis of Primary Endpoint, P030 and P027

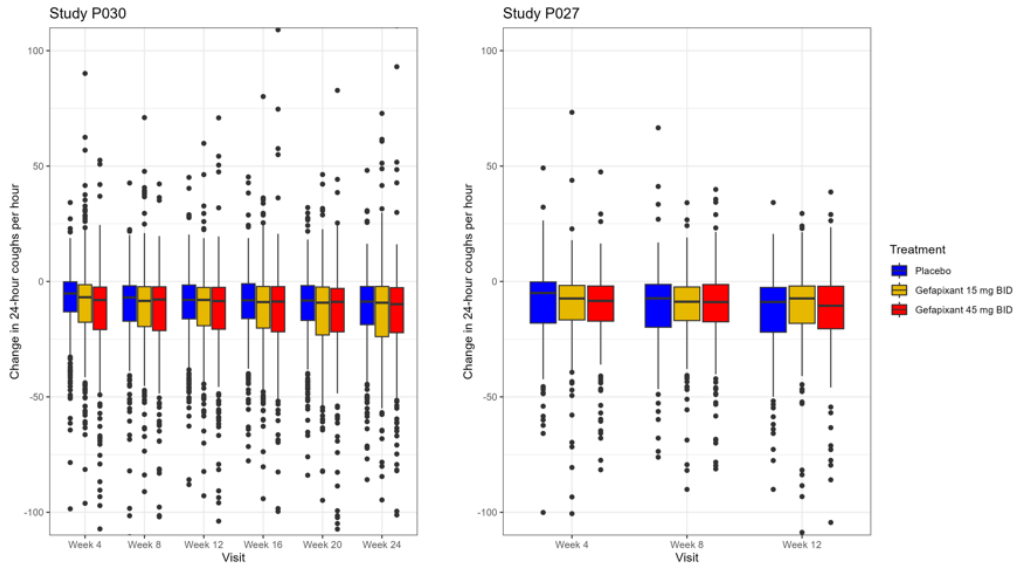
To further understand the magnitude of change from baseline and the amount of treatment separation, several descriptive graphical analyses were conducted. Figure 13 and Figure 14 are boxplots for the studies showing the absolute change in 24-hour cough frequency over time. A closer look at those responses ranging from a reduction of 100 to an increase of 100 is shown in Figure 14 by limiting the y-axis to ±100. These data demonstrate the weak response gefapixant 45 mg dose had on patients and the relatively similar placebo response.

Figure 13. Change in 24-Hour Cough Frequency for Trials P030 and P027 (Full Analysis Set, Recount Data)



Source: adeff.xpt; statistical analyst. See summary statistics in in Table 25 (P030) and Table 26 (P027).
Abbreviation: BID, twice daily

Figure 14. Change in 24-Hour Cough Frequency for Trials P030 and P027 (Y-Axis Restricted to ±100 Cough Frequency, Full Analysis Set, Recount Data, Zoomed In)



Source: adeff.xpt; statistical analyst. See summary statistics in Table 25 (P030) and Table 26 (P027).
Abbreviation: BID, twice daily

Table 25. Change in 24-Hour Cough Frequency for Trial P030

Statistic	Placebo	Gefapixant 15 mg BID	Gefapixant 45 mg BID
Week 4			
N	412	409	403
Mean (SD)	-9.17 (18.04)	-10.01 (18.66)	-13.57 (22.64)
Median	-5.19	-6.83	-8.04
Mode	-0.88	-1.88	-4.08
Minimum	-179.55	-96.11	-140.51
Maximum	34.21	90.15	113.58
Q1, Q3	(-13, 0)	(-18, -1)	(-21, -2)
Week 8			
N	398	389	377
Mean (SD)	-11.52 (18.71)	-11.24 (19.21)	-13.07 (23.11)
Median	-6.92	-8.42	-7.83
Mode	-4.50	-6.71	-26.04
Minimum	-174.13	-91.05	-149.88
Maximum	42.67	112.21	158.13
Q1, Q3	(-17, -2)	(-20, -2)	(-21, -2)
Week 12			
N	383	373	363
Mean (SD)	-11.63 (20.26)	-11.85 (16.85)	-14.01 (26.26)
Median	-7.96	-7.99	-8.50
Mode	-17.79	-5.62	-9.21
Minimum	-181.47	-92.86	-176.18
Maximum	45.10	59.83	183.79
Q1, Q3	(-16, -2)	(-19, -3)	(-21, -3)
Week 16			
N	382	367	355
Mean (SD)	-11.34 (18.81)	-12.38 (18.58)	-14.02 (29.41)
Median	-8.17	-8.92	-8.70
Mode	-6.12	-3.38	-0.38
Minimum	-179.26	-94.10	-183.88
Maximum	45.29	80.12	238.38
Q1, Q3	(-16, -1)	(-20, -2)	(-22, -2)
Week 20			
N	366	362	345
Mean (SD)	-11.63 (21.76)	-13.56 (18.39)	-14.66 (26.14)
Median	-8.23	-9.19	-8.83
Mode	-1.96	-2.88	-5.83
Minimum	-179.84	-94.82	-211.60
Maximum	176.62	46.33	149.82
Q1, Q3	(-17, -2)	(-23, -2)	(-22, -3)
Week 24			
N	368	363	347
Mean (SD)	-13.05 (20.13)	-13.05 (19.96)	-15.68 (26.67)
Median	-8.71	-9.21	-9.83
Mode	-17.25	-3.50	-1.00
Minimum	-180.88	-94.65	-170.93
Maximum	48.12	72.83	110.75
Q1, Q3	(-19, -2)	(-24, -2)	(-22, -3)

Source: adeff.xpt; statistical analyst.

Abbreviations: BID, twice daily; Q, quartile; SD, standard deviation

Table 26. Change in 24-Hour Cough Frequency for Trial P027

Statistic	Placebo	Gefapixant 15 mg BID	Gefapixant 45 mg BID
Week 4			
N	217	224	207
Mean (SD)	-14.36 (51.13)	-10.72 (18.04)	-14.04 (28.51)
Median	-5.00	-7.37	-8.42
Mode	-0.38	-9.50	-5.25
Minimum	-657.08	-100.58	-243.92
Maximum	49.17	73.29	47.46
Q1, Q3	(-18, 0)	(-17, -2)	(-17, -2)
Week 8			
N	208	218	199
Mean (SD)	-17.69 (60.75)	-12.2 (17.98)	-14.2 (31.76)
Median	-7.31	-8.77	-8.92
Mode	-4.21	-10.58	-2.83
Minimum	-792.04	-117.00	-271.92
Maximum	66.58	34.12	39.88
Q1, Q3	(-20, -1)	(-17, -2)	(-17, -1)
Week 12			
N	205	210	194
Mean (SD)	-18.5 (58.28)	-12.31 (18.74)	-15.9 (31.43)
Median	-8.87	-7.35	-10.52
Mode	-2.71	-32.62	-20.88
Minimum	-711.92	-108.71	-277.88
Maximum	162.44	29.46	38.75
Q1, Q3	(-22, -3)	(-18, -2)	(-20, -2)

Source: adeff.xpt; statistical analyst.

Abbreviations: BID, twice daily; Q, quartile; SD, standard deviation

Table .27. Trial P030 Summary Statistics for 24-Hour Cough Frequency

Statistic	Placebo	Gefapixant 15 mg BID	Gefapixant 45 mg BID
Baseline			
N	432	431	434
Mean (SD)	28.51 (24.56)	28.12 (22.17)	28.62 (29.89)
Median	21.25	22.12	19.88
Mode	16.92	3.88	14.96
Minimum	0.71	1.00	0.17
Maximum	183.59	151.62	230.10
Q1, Q3	(12.42, 37.08)	(11.36, 39.54)	(10.88, 37.86)
Week 24			
N	369	370	352
Mean (SD)	15.19 (15.47)	15.41 (19.16)	12.86 (17.65)
Median	11.38	9.46	7.71
Mode	12.42	2.54	4.00
Minimum	0.00	0.00	0.12
Maximum	93.42	213.21	161.29
Q1, Q3	(4.21, 20.83)	(3.76, 20.43)	(3.56, 15.09)

Source: adeff.xpt; statistical analyst.

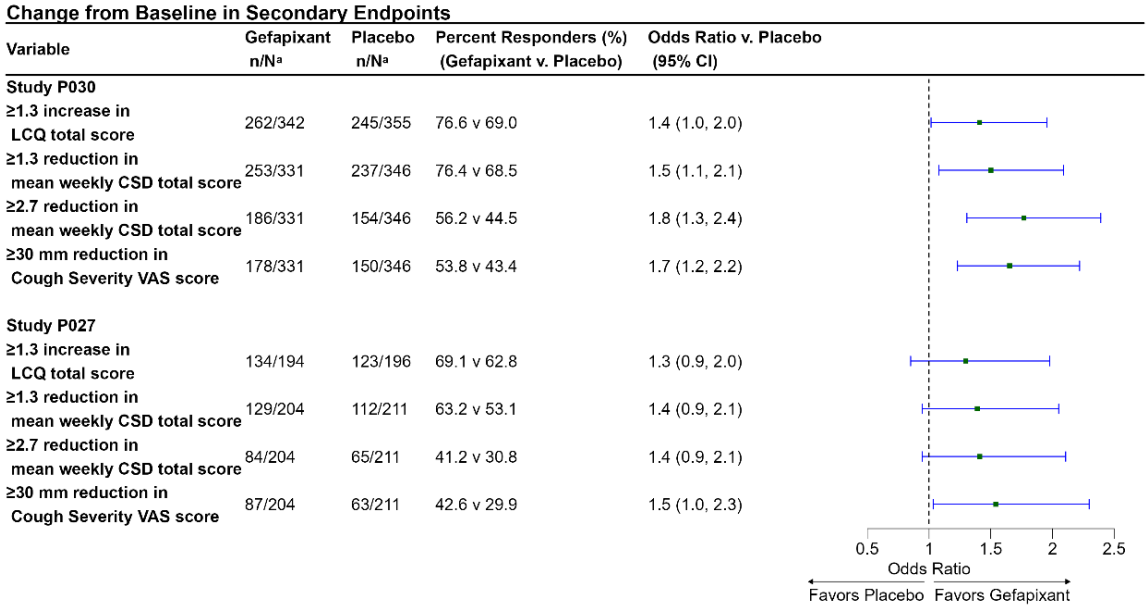
Abbreviations: BID, twice daily; Q, quartile; SD, standard deviation

Table 28. Trial P027 Summary Statistics for 24-Hour Cough Frequency

Statistic	Placebo	Gefapixant 15 mg BID	Gefapixant 45 mg BID
Baseline			
N	232	235	237
Mean (SD)	39.47 (81.06)	28.01 (22.03)	30.22 (39.38)
Median	26.10	21.83	20.88
Mode	9.46	16.71	15.00
Minimum	0.33	0.79	0.17
Maximum	1053.46	132.83	399.12
Q1, Q3	(12.87, 45.5)	(13.17, 37.33)	(12.17, 36.21)
Week 12			
N	205	210	194
Mean (SD)	21.36 (46.29)	16.04 (15.07)	14.9 (19.44)
Median	11.56	10.94	8.67
Mode	7.96	9.83	4.96
Minimum	0.25	0.08	0.04
Maximum	566.49	103.33	175.38
Q1, Q3	(5.2, 26.34)	(5.78, 22.53)	(3.77, 19.15)

Source: adeff.xpt; statistical analyst.
Abbreviations: BID, twice daily; Q, quartile; SD, standard deviation

Figure 15. Odds Ratio in Proportion of Responders for Results of Patient-Reported Outcome-Based Secondary Endpoints

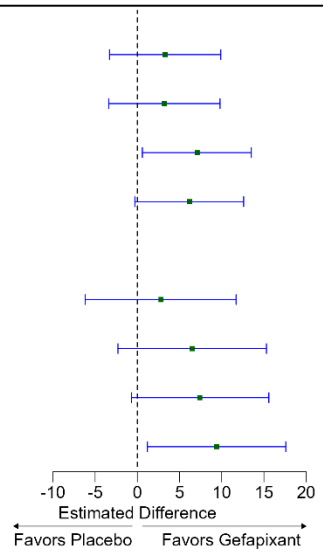


Source: Statistical analyst and Tables 10 and 14.
Gefapixant: gefapixant 45 mg. Change from baseline at Week 24 (Trial P030)/12 (Trial P027).
n, number of responders; N, number of subjects with available data at Week 24 for Trial P030 (Week 12 for Trial P027) with terms of treatment, visit, treatment-by-visit interaction, gender, region, baseline, and baseline-by-visit interaction among subjects who had baseline and post baseline values.
There was a small discrepancy in odds ratio results for ≥ 2.7 CSD reduction for Trial P027.
Abbreviations: CI, confidence interval; CSD, cough severity diary; LCQ, Leicester Cough Questionnaire; VAS, visual analog scale

Figure 16. Difference in Proportion of Responders for Patient-Reported Outcome-Based Secondary Endpoints

Change from Baseline in Secondary Endpoints

Variable	Gefapixant n/N[b]	Placebo n/N[b]	Percent Responders (%) (Gefapixant v. Placebo)	Estimated Difference (%) (95% CI)
Study P030				
≥1.3 increase in LCQ total score	262/419	245/415	62.5 v 59.0	3.3 (-3.3, 9.9)
≥1.3 reduction in mean weekly CSD total score	253/437	237/434	57.9 v 54.6	3.2 (-3.4, 9.8)
≥2.7 reduction in mean weekly CSD total score	186/437	154/434	42.6 v 35.5	7.1 (0.6, 13.5)
≥30 mm reduction in Cough Severity VAS score	178/437	150/434	40.7 v 34.6	6.2 (-0.3, 12.6)
Study P027				
≥1.3 increase in LCQ total score	134/236	123/229	56.8 v 53.7	2.8 (-6.2, 11.7)
≥1.3 reduction in mean weekly CSD total score	129/243	112/241	53.1 v 46.5	6.5 (-2.3, 15.3)
≥2.7 reduction in mean weekly CSD total score	84/243	65/241	34.6 v 27.0	7.4 (-0.7, 15.6)
≥30 mm reduction in Cough Severity VAS score	87/243	63/241	35.8 v 26.1	9.4 (1.2, 17.6)



Source: Statistical analyst and Tables 10 and 14.
Gefapixant, gefapixant 45 mg. Change from baseline at Week 24 (Trial P030) or Week 12 (Trial P027).
n, number of responders; N[b], number of subjects who had baseline values.
Estimated difference is based on the stratified Miettinen and Nurminen method with gender and region.
Abbreviations: CI, confidence interval; CSD, Cough Severity Diary; LCQ, Leicester Cough Questionnaire; VAS, visual analog scale

Effect of Taste Disturbance on Primary Endpoint Results

FDA performed post hoc, exploratory analyses of the potential relationship of taste-related AE incidence and the primary endpoint results in each pivotal trial. For this analysis, we focused on the results in the gefapixant 45mg arm and not the relative reduction compared to placebo, because it was unclear whether the population of placebo patients experiencing taste disturbance was representative of the population. In trials P030 and P027, the model-based geometric mean ratio in the 283 (P030) and 134 (P027) gefapixant 45mg arm patients experiencing taste disturbance was 0.33 (P030) and 0.39 (P027); in the 126 (P030) and 83 (P027) patients in the 45 mg subgroup not experiencing taste disturbance, the mean ratio was 0.44 (P030; Table 29) and 0.44 (P027; Table 30).

These results suggest a potential association between incidence of taste-related AEs and reduction in cough frequency. Conclusions cannot be made regarding the true impact of the taste disturbance on study outcomes, yet these data highlight the uncertainty of the efficacy data in this program, due to the inherent nature of CC and the hypothesized mechanism of action and adverse event profile of this product.

Table 29. Analysis of 24-Hour Cough Frequency, Trial P030 at Week 24 Subgroup Analysis by Whether Subjects Experienced Taste Disturbance (Full Analysis Set, Recount Data, MMRM)

	Subjects Experiencing Taste Disturbance			Subjects Not Experiencing Taste Disturbance		
Statistic	Placebo	Gefapixant 15 mg	Gefapixant 45 mg	Placebo	Gefapixant 15 mg	Gefapixant 45 mg
N	38	86	283	381	329	126
Geometric mean at baseline	21.5	22.6	20.4	20.3	19.6	17.2
Geometric mean at week 24	9.0	10.3	7.0	8.7	7.9	7.5
Geometric mean ratio*	0.42	0.46	0.34	0.43	0.40	0.44
Model based geometric mean ratio (95% CI)**	0.39 (0.28, 0.55)	0.44 (0.35, 0.55)	0.33 (0.29, 0.38)	0.44 (0.39, 0.50)	0.42 (0.37, 0.47)	0.44 (0.36, 0.54)
Relative reduction (95% CI)***		12.4 (-24.3, 66.8)	-16.1 (-41.2, 19.7)		-5.5 (-19.10, 10.37)	-0.6 (-19.72, 23.21)

Source: adeff.xpt; statistical analyst.

Geometric means were used because these frequency data were log-transformed.

* Based on subjects with nonmissing values at baseline and Week 24.

** Based on the MMRM model.

*** The estimated relative reduction (relative to placebo) is calculated by 100 (e**DIFF -1). DIFF is the treatment difference in change from baseline at Week 24 based on the log-transformed data.

Abbreviations: CI, confidence interval; MMRM, mixed model repeated measures; N, number of subjects who had baseline and postbaseline assessments

Table 30. Analysis of 24-Hour Cough Frequency, Trial P027 at Week 12 Subgroup Analysis by Whether Subjects Experienced Taste Disturbance (Full Analysis Set, Recount Data, MMRM)

	Subjects Experiencing Taste Disturbance			Subjects Not Experiencing Taste Disturbance		
Statistic	Placebo	Gefapixant 15 mg	Gefapixant 45 mg	Placebo	Gefapixant 15 mg	Gefapixant 45 mg
N	10	29	134	212	198	83
Geometric mean at baseline	36.8	21.5	22.1	23.1	20.8	14.7
Geometric mean at Week 24	15.0	11.3	7.5	10.5	10.1	7.2
Geometric mean ratio*	0.41	0.52	0.34	0.45	0.48	0.49
Model based geometric mean ratio (95% CI)**	0.53 (0.26, 1.08)	0.61 (0.39, 0.97)	0.39 (0.30, 0.50)	0.47 (0.40, 0.54)	0.47 (0.41, 0.55)	0.44 (0.35, 0.54)
Relative reduction (95% CI)***		15.2 (-46.5, 147.9)	-27.0 (-63.7, 46.6)		1.8 (-16.05, 23.32)	-6.3 (-27.51, 21.13)

Source: adeff.xpt; statistical analyst.

Geometric means were used because these frequency data were log-transformed.

* Based on subjects with nonmissing values at baseline and Week 12.

** Based on the MMRM model.

*** The estimated relative reduction (relative to placebo) is calculated by 100 (e^{DIFF} -1). DIFF is the treatment difference in change from baseline at Week 24 based on the log-transformed data.

Abbreviations: CI, confidence interval; MMRM, mixed model repeated measures; N, number of subjects who had baseline and postbaseline assessments

Cough Count Reduction ≥50% and ≥70%

Outside the multiplicity hierarchy, higher thresholds for reduction in 24-hour cough frequency of 50% and 70% were also evaluated by the Applicant (Table 31). In addition, the FDA graphed the percentage reduction in 24-hour cough frequency plotted against the frequency of patients experiencing that percentage improvement (Figure 6 with corresponding numerical presentations in Table 32 and Table 33).

Table 31. Analysis of Subjects With ≥50% and ≥70% Reductions From Baseline in 24-Hour Cough Frequency at Week 24 (Trial P030) and Week 12 (Trial P027) (Full Analysis Set)

Endpoint	Variable	Placebo	Gefapixant 15 mg	Gefapixant 45 mg
Trial P030	N	419	415	409
≥50% reduction from baseline in 24-hour cough frequency at Week 12	n (%)	185 (44)	194 (47)	202 (49)
	Odds ratio vs. placebo*		1.11 (0.83, 1.48)	1.36 (1.02, 1.83)
≥70% reduction from baseline in 24-hour cough frequency at Week 12	n (%)	112 (27)	116 (28)	134 (33)
	Odds ratio vs. placebo*		1.06 (0.77, 1.44)	1.42 (1.05, 1.94)
Trial P027	N	222	227	217
≥50% reduction from baseline in 24-hour cough frequency at Week 12	n (%)	94 (42)	88 (39)	105 (48)
	Odds ratio vs. placebo*		0.92 (0.62, 1.36)	1.39 (0.94, 2.08)
≥70% reduction from baseline in 24-hour cough frequency at Week 12	n (%)	51 (23)	49 (22)	65 (30)
	Odds ratio vs. placebo*		0.96 (0.61, 1.52)	1.53 (0.99, 2.37)

Source: Applicant's Tables 4-5 and 4-7, CSR Addenda

≥50% and >70% reductions from baseline in 24-hour cough frequency at Week 24 (P030) and Week 12 (P027).

* Based on the logistic regression model among subjects who had baseline and postbaseline values. The covariates include treatment, visit, treatment-by-visit interaction, gender, region, baseline 24-hour cough frequency and the interaction of baseline 24-hour cough frequency.

Abbreviations: CSR, clinical study report; N, number of subjects with available data at Week 24 (P030) and Week 12 (P027); n, number of responders at Week 24 (P030) and Week 12 (P027).

Table 32. Trial P030 Cumulative Responder Cough Reduction

Percentage Reduction in Cough Frequency From Baseline at Week 24	Placebo n (%)	MK-7264 15 mg BID n (%)	MK-7264 45 mg BID n (%)
N	419	415	409
>0%	300 (71.6)	306 (73.7)	300 (73.3)
≥10%	286 (68.3)	289 (69.6)	285 (69.7)
≥20%	269 (64.2)	274 (66)	267 (65.3)
≥30%	245 (58.5)	248 (59.8)	248 (60.6)
≥40%	218 (52)	228 (54.9)	227 (55.5)
≥50%	185 (44.1)	194 (46.8)	202 (49.4)
≥60%	143 (34.1)	158 (38.1)	172 (42)
≥70%	112 (26.7)	116 (28)	134 (32.8)
≥80%	69 (16.5)	69 (16.6)	88 (21.5)
≥90%	25 (6)	33 (8)	39 (9.5)
100%	1 (0.2)	1 (0.2)	0 (0)

Source: adeff.xpt; statistical analyst.

Abbreviations: BID, twice daily; MK-7264, gefapixant

Table 33. Trial P027 Cumulative Responder Cough Reduction

Percentage Reduction in Cough Frequency From Baseline at Week 12	Placebo (%)	MK-7264 15 mg BID (%)	MK-7264 45 mg BID (%)
N	222	227	217
≥0%	172 (77.5)	176 (77.5)	159 (73.3)
≥10%	160 (72.1)	160 (70.5)	152 (70)
≥20%	152 (68.5)	146 (64.3)	145 (66.8)
≥30%	135 (60.8)	135 (59.5)	133 (61.3)
≥40%	118 (53.1)	110 (48.5)	122 (56.2)
≥50%	94 (42.3)	88 (38.8)	105 (48.4)
≥60%	74 (33.3)	70 (30.8)	83 (38.2)
≥70%	51 (23)	49 (21.6)	65 (30)
≥80%	31 (14)	31 (13.7)	43 (19.8)
≥90%	14 (6.3)	10 (4.4)	23 (10.6)
100%	0 (0)	0 (0)	0 (0)

Source: adeff.xpt; statistical analyst.

Abbreviations: BID, twice daily; MK-7264, gefapixant

Table 34. Trial P030 Summary Statistics for Change in 24-Hour Cough Frequency by PGIC at Week 24 (Full Analysis Set, Recount Data)

Statistic	Placebo	Gefapixant 15 mg BID	Gefapixant 45 mg BID
PGIC=1 (Very much improved)			
N	71	80	95
Mean (SD)	-20.5 (21.8)	-21.6 (22.2)	-18.1 (16.1)
Median, mode	-13.2, -24	-19.2, -19	-13.9, -8.5
Min, max	-120.9, 8.9	-84.5, 61.6	-79.3, 4.5
Q1, Q3	-31.7, -6.5	-34.5, -7.5	-25.3, -6.6
PGIC=2 (Much improved)			
N	108	113	127
Mean (SD)	-17.1 (23.2)	-13.2 (17.9)	-18.8 (27.9)
Median, mode	-10.8, -17.3	-8.7, -8.7	-10.9, -5.6
Min, max	-180.9, 10.2	-68.1, 51.3	-170.9, 48.5
Q1, Q3	-20.4, -5.3	-25.0, -1.9	-28.1, -2.9
PGIC=3 (Minimal improved)			
N	91	100	62
Mean (SD)	-8.8 (14.4)	-12.5 (17.9)	-13.5 (39.8)
Median, mode	-7.1, -4.5	-8.1, -21.3	-6.8, -2.3
Min, max	-85.9, 26.2	-94.7, 41.6	-167.1, 110.8
Q1, Q3	-15.7, 0.3	-21.4, -2.1	-20.2, -1.6
PGIC=4 (No change)			
N	75	56	49
Mean (SD)	-7.8 (16.2)	-3.9 (21.1)	-7.7 (19.7)
Median, mode	-3.8, 0.3	-3.3, -3.3	-4.4, -9.8
Min, max	-56.8, 30.3	-64.5, 72.8	-126.1, 16.1
Q1, Q3	-14.4, 0.5	-11.1, 1.5	-10.0, 0.0
PGIC=5 (Minimal worse)			
N	13	7	6
Mean (SD)	-2.8 (15.6)	-1.4, 5.5	-14.1 (13.5)
Median, mode	-2.4, .	-2.8, .	-11.8, .
Min, max	-25.6, 30.7	-7.3, 8.29	-38.9, -1.0
Q1, Q3	-12.7, 5.8	-5.0, 0.8	-16.1, -5.8
PGIC=6 (Much worse)			
N	4	3	5
Mean (SD)	5.7 (32.6)	-2.9 (14.5)	-4.3 (15)
Median, mode	2.6, .	1, .	-1, .
Min, max	-30.4, 48.1	-18.9, 9.4	-22.1, 14.7
Q1, Q3	-9.7, 18.0	-9.0, 5.2	-16.8, 3.5
PGIC=7 (Very much worse)			
N	None	1	1
Mean (SD)	-	3.9 (.)	-1 (.)
Median, mode	-	3.9, 3.9	-1, -1
Min, max	-	3.9, 3.9	-1, -1
Q1, Q3	-	3.9, 3.9	-1, -1

Source: adeff.xpt; statistical analyst.

Mode is not reported if no repetitions occur in the data.

Abbreviations: BID, twice daily; N, none; no observation; PGIC, patient global impression of change; Q, quartile; SD, standard deviation

Table 35. Trial P027 Summary Statistics for Change in 24-Hour Cough Frequency by PGIC at Week 12 (Full Analysis Set, Recount Data)

Statistic	Placebo	Gefapixant 15 mg BID	Gefapixant 45 mg BID
PGIC=1 (Very much improved)			
N	20	32	39
Mean (SD)	-27.5 (22.6)	-17.9 (21.5)	-29.8 (27.7)
Median, mode	-22.2, .	-14.1, .	-17.5, -20.9
Min, max	-72.8, -1.9	-93.2, 24	-119.9, -1.2
Q1, Q3	-44.6, -8.1	-28.5, -6.5	-41.5, -11.9
PGIC=2 (Much improved)			
N	47	52	57
Mean (SD)	-18.6 (23.8)	-15.9 (16.1)	-14.4 (29.7)
Median, mode	-12.2, .	-14.3, -32.6	-9.9, .
Min, max	-147.8, 6.5	-81.7, 16.7	-215.7, 22.8
Q1, Q3	-24.9, -5.6	-26.9, -4.7	-20.5, -3.7
PGIC=3 (Minimal improved)			
N	59	67	53
Mean (SD)	-32.5 (100.8)	-10.6 (16.6)	-12.3 (18.9)
Median, mode	-5.5, .	-6.7, -2.2	-7.8, .
Min, max	-711.9, 13.3	-83.8, 23	-104.4, 17
Q1, Q3	-23.0, -2.3	-18.2, -2.4	-16.7, -1.7
PGIC=4 (No change)			
N	61	41	30
Mean (SD)	-8.3 (12.9)	-2.6 (8.9)	-12.1 (53)
Median, mode	-8.3, -12.3	-1.9, .	-5.4, .
Min, max	-41.6, 34.2	-23.7,29.5	-277.9, 38.8
Q1, Q3	-13.7, -1.9	-6.5, 1.0	-13.2, 1.0
PGIC=5 (Minimal worse)			
N	6	3	4
Mean (SD)	4.1 (10.8)	-26.3 (27)	-0.9 (3.4)
Median, mode	1.1, .	-27.1, .	-1.2, .
Min, max	-8.7, 20.6	-52.9, 1.1	-4.6, 3.3
Q1, Q3	-2.2, 10.3	-40.0, -13.0	-3.0, 1.0
PGIC=6 (Much worse)			
N	4	1	3
Mean (SD)	32.8 (86.9)	-8.8 (.)	-2.4 (15.7)
Median, mode	-7.2, .	-8.8, -8.8	0.1, .
Min, max	-17, 162.4	-8.8, -8.8	-19.1, 11.9
Q1, Q3	-16.0, 41.6	-8.8, -8.8	-9.5, 6.0
PGIC=7 (Very much worse)			
N	1	3	None
Mean (SD)	-25.5 (.)	-12.1 (20.1)	-
Median, mode	-25.5, -25.5	-7.6, .	-
Min, max	-25.5, -25.5	-34, 5.3	-
Q1, Q3	-25.5, -25.5	-20.8, -1.2	-

Source: adeff.xpt; statistical analyst.

Mode is not reported if no repetitions occur in the data.

Abbreviations: BID, twice daily; N, none; no observation; PGIC, patient global impression of change; Q, quartile; SD, standard deviation

6.2 Supplemental Trial P042: Additional Information

6.2.1 Brief Protocol Summary

This trial is described for completeness, but due to the inherent study design limitations outlined below, the results do not aid in the clinical interpretation of efficacy for gefapixant. The Applicant did not request FDA feedback on the study design or its potential to support efficacy and labeling claims until after trial completion.

Title: **A Phase 3b, Randomized, Double-Blind, Placebo-Controlled, Multicenter Study to Evaluate the Efficacy and Safety of Gefapixant in Women with Chronic Cough and Stress Urinary Incontinence**

Dates

May 10, 2020, to September 2, 2022; CSR completed January 25, 2023

Trial Design

This was a 12-week, multicenter, randomized, double-blind, placebo-controlled, parallel-group trial to assess the efficacy and safety of gefapixant in reducing the frequency of cough-induced SUI in female subjects at least 18 years of age with RCC or UCC. The trial enrollment criteria relevant to the diagnosis of RCC or UCC were well-matched to those in the pivotal trials (P030 and P027). The enrollment criteria are intended to define a population with "cough-induced" SUI, but this is not recognized as a distinct subpopulation given that SUI has multiple triggers, such as sneezing, laughing, and other activities causing increased abdominal pressure. Likewise, FDA does not consider "cough-induced" SUI to be a standalone indication.

After screening and a single-blind, placebo run-in period of two weeks, 376 subjects were randomized 1:1 to gefapixant 45 mg or placebo taken orally by tablet twice daily. Subjects continued study intervention throughout the Treatment Period for 12 weeks, at which time efficacy was assessed, and had a safety follow-up telephone call 14 days after completion of study intervention. Assessments for efficacy included collection of various PROs, including a diary to collect episodes of incontinence. This trial did not record or analyze coughs to provide an assessment of cough frequency.

Endpoints

- Primary endpoint: episodes of cough-induced SUI, defined as percent change in average daily cough-induced SUI episodes from baseline at Week 12
- Secondary efficacy endpoints: none

Clinical trials for SUI are expected to evaluate the change in all-cause (not cough-induced) incontinence episodes as a coprimary endpoint alongside a fit-for-purpose PRO. In this trial, no secondary efficacy endpoints were assessed to facilitate the clinical interpretation of the primary endpoint.

6.2.2 Subject Disposition

Subject disposition in trial P042 is summarized in Table 36. Similar to what was observed in the pivotal trials of gefapixant for CC, dose-dependent treatment discontinuation and study discontinuation occurred, and the resultant missing data may impact the efficacy assessment.

Table 36. P042 Disposition of Subjects (All Subjects Randomized)

Variable	Placebo n (%)	Gefapixant 45 mg BID n (%)	Total n (%)
Participants in population	190	186	376
Participants treated	190	185	375
Study status			
Started	190	186	376
Completed	184 (96.8)	176 (94.6)	360 (95.7)
Discontinued	6 (3.2)	10 (5.4)	16 (4.3)
Withdrawal by subject	4 (2.1)	8 (4.3)	12 (3.2)
Other	2 (1.1)	2 (1.1)	4 (1.1)
Treatment status for trial			
Started	190	185	375
Completed	182 (95.8)	166 (89.7)	348 (92.8)
Discontinued	8 (4.2)	19 (10.3)	27 (7.2)
Adverse event[1]	2 (1.1)	13 (7.0)	15 (4.0)
Withdrawal by subject	4 (2.1)	5 (2.7)	9 (2.4)
Other	2 (1.1)	1 (0.5)	3 (0.8)

Source: Clinical study report.

Treatment arms are based on the randomization allocation at baseline regardless of the actual treatment received.

[1] Participants reported as discontinued treatment due to one or more adverse events.

Abbreviation: BID, twice daily

6.2.3 Baseline Demographic and Disease Characteristics

Baseline subject demographic and disease characteristics for P042 are shown in Table 37. Of the randomized subjects who received at least one dose of study drug, they were all female (consistent with the inclusion criteria), majority White, and had a mean age of 56 years. Subject demographic and disease characteristics were generally similar between treatment arms. Overall, there are no observed differences in demographic and disease characteristics across arms that would be expected to impact the evaluation of efficacy.

Table 37. P042 Subject Baseline Characteristics (All Subjects Randomized and Treated)

Variable	Placebo N=190	Gefapixant 45 mg BID N=185	Total N=375
Sex, n (%)			
Female	190 (100)	185 (100)	375 (100)
Age (years), n (%)			
<60 years	106 (55.8)	110 (59.5)	216 (57.6)
≥60 years	84 (44.2)	75 (40.5)	159 (42.4)
Age (years)			
Mean (SD)	56.6 (11.3)	56.2 (11.6)	56.4 (11.4)
Median	56.5	57.0	57.0
Min, max	25, 83	22, 82	22, 83
Race, n (%)			
American Indian or Alaska Native	18 (9.5)	16 (8.6)	34 (9.1)
Asian	3 (1.6)	6 (3.2)	9 (2.4)
Black or African American	0 (0.0)	3 (1.6)	3 (0.8)
Multiple	26 (13.7)	24 (13.0)	50 (13.3)
White	143 (75.3)	136 (73.5)	279 (74.4)

Variable	Placebo N=190	Gefapixant 45 mg BID N=185	Total N=375
Ethnicity, n (%)			
Hispanic or Latino	79 (41.6)	73 (39.5)	152 (40.5)
Not Hispanic or Latino	111 (58.4)	112 (60.5)	223 (59.5)
Region, n (%)			
North America	11 (5.8)	9 (4.9)	20 (5.3)
Europe	102 (53.7)	101 (54.6)	203 (54.1)
Asia Pacific	3 (1.6)	3 (1.6)	6 (1.6)
Other	74 (38.9)	72 (38.9)	146 (38.9)
Primary diagnosis, n (%)			
Refractory chronic cough	149 (78.4)	140 (75.7)	289 (77.1)
Unexplained chronic cough	41 (21.6)	45 (24.3)	86 (22.9)
Baseline mean weekly Cough Severity VAS (mm), n (%)			
<60 mm	51 (26.8)	55 (29.7)	106 (28.3)
≥60 mm	139 (73.2)	130 (70.3)	269 (71.7)
Baseline mean weekly Cough Severity VAS (mm)			
Participants with data	190	185	375
Mean (SD)	69.5 (15.6)	69.3 (15.8)	69.4 (15.7)
Median	71.1	67.7	69.5
Min, max	28.4, 100.0	22.6, 100.0	22.6, 100.0
Duration of chronic cough (months)			
Participants with data	143	154	297
Mean (SD)	5.1 (6.6)	5.2 (6.5)	5.2 (6.6)
Median	2.9	2.7	2.8
Min, max	1.2, 41.1	1.2 to 40.0	1.2 to 41.1
Duration of stress urinary incontinence (months)[1]			
Participants with data	190	185	375
Mean	53.8	42.9	48.4
SD	80.8	58.0	70.6
Median	28.2	24.6	26.1
Min, max	3.9, 586.6	1.6, 458.4	1.6, 586.6
Baseline mean daily cough-induced stress urinary incontinence episodes, 7-day average			
Less than the median (3.86)	96 (50.5)	89 (48.1)	185 (49.3)
Greater than or equal to the median (3.86)	93 (48.9)	96 (51.9)	189 (50.4)
Missing	1 (0.5)	0 (0.0)	1 (0.3)
Participants with data	189	185	374
Mean	4.7	4.7	4.7
SD	4.1	3.0	3.6
Median	3.7	3.9	3.9
Min, max	2.0, 49.7	1.1, 24.0	1.1, 49.7

Source: Clinical study report.

[1] Calculated as time between date of diagnosis and randomization.

Abbreviations: BID, twice daily; max, maximum; min, minimum; SD, standard deviation; VAS, visual analog scale

6.2.4 Primary Endpoint Results

Week 12 Cough-Induced SUI Episodes (Percentage Change)

Results for the primary endpoint analysis are shown in Table 38.

Table 38. P042 Mean Percentage Change for Cough-Induced SUI Episodes (FAS, MMRM)

Visit	Treatment	N	Mean (SD)	Mean Percentage Change (SD)	Model-Based LS Means (95% CI)	Estimated Difference (95% CI)	P-value
Baseline	Placebo	189	4.7 (4.10)				
	MK-7264 45 mg	185	4.7 (2.99)				
Week 4	Placebo	183	3.5 (3.54)	-23.5 (37.90)	-23.3 (-28.9, -17.8)		
	MK-7264 45 mg	177	3.1 (2.80)	-33.4 (38.05)	-33.6 (-39.2, -28)	-10.23 (-18.06, -2.41)	0.0105
Week 8	Placebo	178	3.1 (2.34)	-32.0 (38.35)	-32.3 (-37.9, -26.8)		
	MK-7264 45 mg	176	2.5 (2.50)	-45.4 (39.00)	-45.7 (-51.3, -40)	-13.3 (-21.2, -5.38)	0.0011
Week 12	Placebo	176	2.6 (2.24)	-41.6 (39.25)	-41.1 (-46.7, -35.4)		
	MK-7264 45 mg	174	2.2 (2.53)	-52.8 (37.02)	-52.8 (-58.4, -47.1)	-11.67 (-19.67, -3.67)	0.0044

Source: adeff.xpt, validated by the statistical analyst.
Model used is the mixed model for repeated measures with treatment, baseline mean daily cough-induced SUI episodes or mean weekly overall incontinence episodes, assessment time point, BMI group (<30, ≥30), and time-point-by-treatment interaction as fixed effects with unstructured covariance structure. The response is percent change from baseline in mean daily cough or weekly overall incontinence episodes at each postbaseline visit up to Week 12.
Abbreviations: BMI, body mass index; CI, confidence interval; FAS, full analysis set; LS, least squares; MK-7264, gefapixant; SD, standard deviation; SUI, stress urinary incontinence

The purpose of the trial was to provide evidence that treatment with gefapixant would improve cough, and in turn cough-induced SUI episodes. The FDA identified several concerns with the selected endpoints and their analysis methods. Because SUI episodes can be triggered by multiple events not limited to cough, cough-induced SUI is not recognized as a standalone indication or definition for an endpoint. Clinical trials for SUI are expected to evaluate the change in all-cause incontinence episodes as a coprimary endpoint alongside a fit-for-purpose PRO. Another important issue with the primary endpoint is that treatment comparison defined by percent change from baseline as an outcome measure has undesirable statistical properties, including its sensitivity to influence by the magnitude of the baseline value. For this reason, clinical trials for SUI typically evaluate the change in absolute number of daily episodes of incontinence (based on the 7-day average of daily episodes), along with a fit-for-purpose PRO as coprimary endpoints, to demonstrate that the reduction in SUI episodes is clinically meaningful. In this trial, however, no secondary efficacy endpoints were assessed to aid in efficacy interpretation.

6.2.5 Secondary Endpoint Results

There are no prespecified secondary efficacy endpoints in the clinical study protocol.

6.2.6 Additional Analysis

Absolute Change in Cough-Induced or All-Cause SUI Episodes

Due to the undesirable statistical properties when evaluating percent change from baseline as an outcome measure, the FDA also analyzed absolute change in the average daily cough-induced SUI episodes and all-cause incontinence episodes. The mean absolute change in daily cough-induced SUI episodes at Week 12 was -2.1 in the placebo group and -2.5 in the gefapixant group with a treatment difference of -0.4 episodes (95% CI: -0.9 to 0.1, nominal p=0.093). The mean absolute change in all-cause incontinence episodes at Week 12 was -2.3 in the placebo group and -2.7 in the gefapixant group with a treatment difference of -0.5 (95% CI: -1.1 to 0.2, nominal p=0.145). The Applicant did not include these absolute change endpoints in their study report. The results of these post hoc analyses of absolute

change in SUI episodes, either induced by all triggers or by cough only, do not support a meaningful treatment effect from gefapixant, nor are they nominally statistically significant. As such, the absolute change results do not offer clinically or statistically convincing evidence of efficacy, and no secondary endpoints were assessed in this trial.

6.3 Supplementary Trial P043: Additional Information

6.3.1 Brief Protocol Summary

The major design differences between the pivotal trials and P043 are the population (recent-onset RCC or UCC, i.e., shorter duration of chronic cough), the shorter trial duration (12 weeks of treatment), and the primary endpoint (LCQ total score rather than 24-hour cough frequency). We have included this trial and results for completeness, but we do not believe this trial provides substantive additional information regarding the clinical benefit of gefapixant. This trial does not assess cough frequency, which is the primary basis for establishing effectiveness of gefapixant. This trial assesses the LCQ total score as the primary endpoint. As discussed in the body of the briefing document, there are limitations and uncertainties with the PROs, including LCQ. We also note that the results of this trial show a numerically small treatment difference on the LCQ total score between gefapixant 45 mg and placebo, which is consistent with the small treatment effect seen in the pivotal trials.

Title: A Phase 3b Randomized, Double-blind, Placebo Controlled, Multicenter Study to Evaluate the Efficacy and Safety of Gefapixant in Adult Participants with Recent-Onset Chronic Cough

Dates

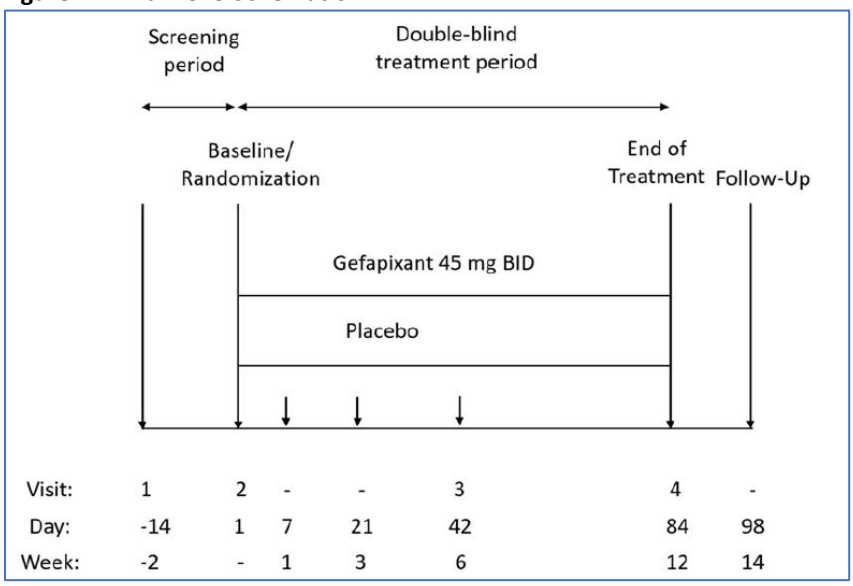
May 21, 2020, to November 3, 2021; CSR completed March 24, 2022

Trial Design

This was a 12-week, multicenter, randomized, double-blind, placebo-controlled, parallel-group trial to assess the efficacy and safety of gefapixant in subjects who had recent onset (<12 months) of either RCC or UCC. Overall, the trial population is well-matched to the population in P030 and P027, except that P043 subjects have a shorter duration of CC (recent-onset CC). This subset of the CC population would not necessarily imply a distinct indication for use of gefapixant, and it is conceptually reasonable that data from a trial in these subjects could contribute support for the broader CC indication.

After screening, 419 subjects were randomized 1:1 to gefapixant 45 mg or placebo taken orally by tablet twice daily. Of the 419 randomized, 415 were treated: 209 with placebo and 206 with gefapixant. Subjects continued study intervention throughout the Treatment Period for 12 weeks, at which time efficacy was assessed, and had a safety follow-up telephone call 14 days after completion of study intervention. Assessments for efficacy included collection of various PROs. This study did not record or analyze coughs to provide an assessment of cough frequency. The trial schematic is shown in Figure 17.

Figure 17. Trial P043 Schematic



Source: Applicant's clinical study protocol, Fig. 1, p. 347.
Abbreviation: BID, twice daily

Endpoints

- Primary endpoint: change from baseline in the LCQ total score at Week 12
- Secondary efficacy endpoint: change from baseline in the Cough Severity VAS score at Week 12

The primary endpoint was analyzed by a MMRM model including treatment arm, visit, interaction of treatment by visit, gender, and baseline LCQ total score.

6.3.2 Subject Disposition

Subject disposition in Trial P043 is summarized in Table 39. Similar to what was observed in the pivotal trials of gefapixant for CC, dose-dependent treatment discontinuation and study discontinuation occurred.

Table 39. P043 Disposition of Subjects (All Subjects Randomized)

Variable	Placebo n (%)	Gefapixant 45 mg BID n (%)	Total n (%)
Participants in population	211	208	419
Participants treated	209 (99.1)	206 (99)	415 (99)
Study status			
Started	211	208	419
Completed	201 (95.3)	192 (92.3)	393 (93.8)
Discontinued	10 (4.8)	16 (7.7)	26 (6.2)
Death	1 (0.5)	0 (0.0)	1 (0.2)
Physician decision	1 (0.5)	1 (0.5)	2 (0.5)
Withdrawal by subject	3 (1.4)	10 (4.8)	13 (3.1)
Other	5 (2.4)	5 (2.4)	10 (2.4)
Treatment status for trial			
Started	209	206	415
Completed	198 (93.8)	175 (84.1)	373 (88.9)
Discontinued	11 (5.3)	31 (15.1)	42 (10.2)
Adverse event[a]	4 (1.9)	23 (11.2)	27 (6.5)
Death	1 (0.5)	0 (0.0)	1 (0.2)
Lost to follow-up[b]	1 (0.5)	0 (0.0)	1 (0.2)
Withdrawal by subject	3 (1.4)	6 (2.9)	9 (2.2)
Other	2 (1.0)	2 (1.0)	4 (1.0)

Source: adbase.xpt; validated by the statistical analyst.

Treatment arms are based on the randomization allocation at baseline regardless of the actual treatment received.

One subject in the placebo group was reported as lost to follow-up for treatment status, however, reason for discontinuing treatment early was withdrawal by subject.

[a] Participants reported as discontinued treatment due to one or more adverse events.

[b] Reported in database as lost to follow-up in error; actual reason for discontinuing treatment was withdrawal by subject.

Abbreviation: BID, twice daily

6.3.3 Baseline Demographic and Disease Characteristics

Baseline subject demographic and disease characteristics for P043 are shown in Table 40. Of the randomized subjects who received at least one dose of study drug, they were majority female, majority White, and had a mean age of 50 years. The study population was younger than in the pivotal trials, likely due to a different inclusion criterion for duration of CC diagnosis. Subject demographic and disease characteristics were generally similar between treatment arms. Overall, there are no observed differences in demographic and disease characteristics across arms that would be expected to impact the evaluation of efficacy.

Table 40. P043 Subject Baseline Characteristics (All Subjects Randomized and Treated)

Characteristic	Placebo N=209	Gefapixant 45 mg BID N=206	Total N=415
Sex, n (%)			
Female	134 (64.1)	134 (65.0)	268 (64.6)
Male	75 (35.9)	72 (35.0)	147 (35.4)
Age (years), n (%)			
<60 years	135 (64.6)	130 (63.1)	265 (63.9)
≥60 years	74 (35.4)	76 (36.9)	150 (36.1)

Characteristic	Placebo N=209	Gefapixant 45 mg BID N=206	Total N=415
Age (years)			
Mean (SD)	52.5 (13.78)	52.5 (13.78)	52.5 (13.77)
Median	55.0	54.0	55.0
Min, max	18.0, 83.0	18.0, 81.0	18.0, 83.0
Race, n (%)			
American Indian or Alaska Native	27 (12.9)	22 (10.7)	49 (11.8)
Asian	2 (1.0)	3 (1.5)	5 (1.2)
Black or African American	0	3 (1.5)	3 (<1)
White	151 (72.2)	149 (72.3)	300 (72.3)
Multiple	29 (13.9)	29 (14.1)	58 (14.0)
Ethnicity, n (%)			
Hispanic or Latino	75 (35.9)	71 (34.5)	146 (35.2)
Not Hispanic or Latino	133 (63.6)	134 (65.0)	267 (64.3)
Unknown	1 (<1)	1 (<1)	2 (<1)
Region, n (%)			
Asia-Pacific	1 (<1)	2 (1.0)	3 (<1)
Europe	123 (58.9)	122 (59.2)	245 (59.0)
North America	11 (5.3)	10 (4.9)	21 (5.1)
Other	74 (35.4)	72 (35.0)	146 (35.2)
Primary diagnosis, n (%)			
Refractory chronic cough	144 (68.9)	150 (72.8)	294 (70.8)
Unexplained chronic cough	65 (31.1)	56 (27.2)	121 (29.2)
Baseline mean weekly Cough Severity VAS (mm), n (%)			
<60 mm	80 (38.3)	73 (35.4)	153 (36.9)
≥60 mm	129 (61.7)	133 (64.6)	262 (63.1)
Baseline mean weekly Cough Severity VAS (mm)			
Mean (SD)	66.2 (14.9)	67.2 (14.9)	66.7 (14.9)
Median	64.7	65.8	65.4
Min, max	22.3 to 100.0	27.7 to 100.0	22.3 to 100.0
Duration of chronic cough (months)			
Mean (SD)	7.2 (2.68)	7.3 (2.79)	7.2 (2.73)
Median	7.0	8.0	7.5
Min, max	2.0, 12.0	1.0, 12.0	1.0, 12.0
Missing	0	1	1

Source: adsl.xpt, adbase.xp, validated by the statistical analyst.

Abbreviations: BID, twice daily; max, maximum; min, minimum; SD, standard deviation; VAS, visual analog scale

6.3.4 Primary Endpoint Results

Results for the primary endpoint analysis are shown in Table 41. The estimated difference in model-based mean change in LCQ total score from baseline at Week 12 for gefapixant 45 mg BID compared to placebo is 0.75 (95% CI: 0.06, 1.44; p=0.034). For context, the range of possible scores for LCQ total score is 3 to 21, with higher scores representing improvement.

Table 41. P043 LCQ Total Score at Week 12 (Full Analysis Set)

Treatment	N	Mean (SD) at Baseline[1]	Mean (SD) at Week 12[1]	Mean (SD) Change From Baseline (Week 12−Baseline)[1]	Model[2] Based Mean Change From Baseline (Week 12−Baseline) (95% CI)
Placebo	199	11.30 (2.80)	14.73 (3.48)	3.43 (3.74)	3.59 (3.09, 4.09)
Gefapixant	199	10.82 (3.08)	15.32 (3.91)	4.49 (3.91)	4.34 (3.84, 4.83)
Treatment Difference				**Estimated Difference[3] (95% CI)**	**p-Value**
Gefapixant vs. placebo				0.75 (0.06, 1.44)	0.034

Source: P043 CSR, Table 14.2-1.

[1] Based on participants with nonmissing values at both baseline and Week 12.

[2] Based on the longitudinal analysis of covariance model consisting of the change from baseline in LCQ total score at each postbaseline visit (up to Week 12) as response. The model includes terms for treatment group (MK-7264 45 mg and placebo), visit (Weeks 6 and 12), the interaction of treatment by visit, gender, and the baseline LCQ total score. The unstructured covariance matrix is used to model the correlation among repeated measurements.

[3] The estimated difference is the treatment difference in model based mean change from baseline at Week 12.

Abbreviations: CI confidence interval; CSR, clinical study report; MK-7264, gefapixant; N, number of participants included in the analysis.; SD, standard deviation; LCQ, Leicester Cough Questionnaire

In contrast to the pivotal trials, which compared the proportion of responders in the gefapixant arm to placebo, in Trial P043, the analysis of LCQ total score was the estimated treatment difference between gefapixant and placebo in mean change from baseline. The interpretation of this statistically significant yet numerically small change in LCQ total score is challenging. We note that a similar value for the post hoc analysis of the change from baseline in LCQ total score was observed in P030. Similar to the pivotal trials, the placebo response was large. The Applicant did not provide adequate information to determine what change in LCQ total score would correspond to a meaningful change from the patient perspective; however, taken at face value, a <1 point treatment difference on a total score with a possible range of 3 to 21 has questionable clinical significance. Additionally, FDA has concerns that the LCQ total score may not be fit-for-purpose to inform regulatory decisions, as discussed in Section 3.1.3.5.2. Finally, cough frequency data were not captured in this trial; thus, an objective assessment of efficacy (via cough frequency) in addition to the PRO-based primary endpoint cannot be performed.

6.3.5 Secondary Endpoint Results

Results for the prespecified secondary endpoint analysis, the estimated relative reduction in model-based mean change in Cough Severity VAS score from baseline at Week 12 for gefapixant 45 mg BID compared to placebo, are shown in Table 42. For context, the VAS score ranges from 0 to 100. The challenges in use and interpretation of the clinical meaningfulness of the Cough Severity VAS score are discussed in Section 3.1.3.5.4.

Similar to the P043 primary endpoint, the selected analysis of this PRO differs from the analysis conducted in the pivotal trials. In P043, the analysis of the Cough Severity VAS score is the estimated relative reduction of gefapixant compared to placebo in mean change from baseline. Even though the Applicant did not provide adequate information to determine what change would correspond to a meaningful change from the patient perspective, the treatment difference from placebo is small (6.9 points) in relationship to the range of possible scores (0-100) and has unclear clinical significance. As such, the interpretation of a small change in Cough Severity VAS score is challenging.

Table 42. P043 Cough Severity VAS at Week 12 (Full Analysis Set, MMRM)

Treatment	N	Mean (SD) at Baseline*	Mean (SD) at Week 12*	Mean (SD) Change From Baseline (Week 12−Baseline)*	Model Based Mean Change From Baseline (Week 12−Baseline) (95% CI)**
Placebo	198	66.5 (15.09)	40.9 (25.49)	-25.5 (26.67)	-24.9 (-28.4, -21.3)
MK-7264 45 mg BID	191	67.8 (14.72)	34.7 (26.69)	-33.1 (26.27)	-31.8 (-35.4, -28.2)
Treatment Difference			**Estimated Relative Reduction (%) and (95% CI)*** **		**P-value**
MK-7264 45 mg vs. Placebo				-6.9 (-11.9, -2.0)	0.006

Source: adqs.xpt; validated by the statistical analyst.

* Based on participants with nonmissing values at both baseline and Week 12.

** Based on the MMRM model (or longitudinal analysis of covariance model) consisting of the change from baseline in mean weekly Cough Severity VAS score at each postbaseline visit (up to Week 12) as response. The model includes terms for treatment group (MK-7264 45 mg and placebo), visit (Weeks 6 and 12), the interaction of treatment by visit, gender, and the baseline mean weekly Cough Severity VAS score. The unstructured covariance matrix is used to model the correlation among repeated measurements.

*** The estimated difference is the treatment difference in model based mean change from baseline at Week 12.

Abbreviations: BID, twice daily; CI confidence interval; MK-7264, gefapixant; N, number of participants included in the analysis; SD, standard deviation

6.3.6 Additional Analyses

As for the pivotal trials, the FDA conducted a post hoc exploratory analysis on the LCQ physical domain, which may be more relevant to inform regulatory decisions, as discussed previously. Table 43 provides the results for the estimated relative reduction in mean change in LCQ physical domain score from baseline at Week 12 for gefapixant 45 mg BID compared to placebo. The result is statistically significant at 0.23 points (95% CI: 0.03, 0.43; p=0.027), but numerically small given the range of possible scores for the LCQ physical domain (1 to 7) The clinical interpretation of this finding is uncertain. As noted in Section 3.1.3.6, the threshold to define a meaningful within-patient change in the LCQ physical domain is not established and would require supportive analysis, as discussed in Section 6.4.1.

Table 43. P043 LCQ Physical Domain Score at Week 12 (Full Analysis Set)

Treatment	N	Mean (SD) at Baseline*	Mean (SD) at Week 12*	Mean Difference (Week 12−Baseline)*	Model Based Mean Difference (Week 12−Baseline) (95% CI)**
Placebo	193	4.15 (0.89)	5.05 (1.05)	0.90 (1.17)	0.94 (0.80, 1.08)
Gefapixant 45 mg BID	193	4.05 (1.05)	5.24 (1.14)	1.19 (1.11)	1.17 (1.02, 1.31)
Treatment Difference			**Estimated Relative Reduction (%) and (95% CI)*** **		**P-value**
Gefapixant 45 mg vs. placebo				0.23 (0.03, 0.43)	0.027

Source: adqs.xpt; validated Applicant's results in an Information Request (p. 3).

* Based on participants with nonmissing values at both baseline and Week 12.

** Based on the longitudinal analysis of covariance model consisting of the change from baseline in LCQ total score at each postbaseline visit (up to Week 12) as response. The model includes terms for treatment group (MK-7264 45 mg and Placebo), visit (Weeks 6 and 12), the interaction of treatment by visit, gender, and the baseline LCQ total score. The unstructured covariance matrix is used to model the correlation among repeated measurements.

*** The estimated difference is the treatment difference in model based mean change from baseline at Week 12.

Abbreviations: CI confidence interval; CSR, clinical study report; LCQ, Leicester Cough Questionnaire; N, number of participants included in the analysis; SD, standard deviation

6.4 Supplemental PRO-Related Information

In the following sections, we provide an overview of the FDA's general expectations for development of PROs and PRO-based endpoints as outlined in guidances for industry. This is followed by considerations

for the specific PROs evaluated in the gefapixant program and a copy of the LCQ, CSD, Cough Severity VAS, and PGIC tools administered in the pivotal trials.

6.4.1 Development of Patient-Reported Outcome-Based Endpoints

It must be acknowledged that the treatment of CC is a novel indication and that the gefapixant program is one of the first clinical development programs for this indication. In designing the pivotal trials of gefapixant for CC, it was unknown which PRO(s) and corresponding PRO-based endpoint(s) would be most appropriate to contribute to the regulatory evaluation of clinical benefit. Given this lack of precedence, the Applicant proposed to incorporate multiple PROs into secondary and exploratory endpoints. In this context, the FDA agreed to the assessment of a variety of PROs, provided insight on known limitations or concerns about the PROs that would need to be addressed, and advised the Applicant that interpretation of these PRO-based endpoints would be an important element of a future NDA review.

The use of methodologically sound and fit-for-purpose[5] data collection tool(s) to collect patient experience data can provide direct evidence about the clinical benefits and risk of a proposed therapy. Development of a fit-for-purpose tool begins with understanding the target disease or condition (e.g., via natural history of the disease, patient and caregiver perspectives, clinical expert input) and what aspect of the disease or condition the drug is expected to improve in the target patient population. As a general guiding principle, FDA values the input from patients to inform drug development and regulatory decision-making; patients are the experts in the experience of their disease, and they are the ultimate stakeholders in the outcomes of medical treatment. In developing a PRO, insight gleaned from detailed interviews with patients to understand disease burden and experience, the most troublesome symptoms, and desired outcomes of treatment, is essential.

Based on this fundamental knowledge of the target disease or condition and the expected effects of the treatment under study, a critical next step in the selection or development of PROs is identifying the appropriate concept(s) of interest[6] for the target context of use. In clinical trials, concepts measured by a PRO should: (1) be sufficiently prevalent in the target patient population and reflect what is important to patients; (2) demonstrate change or stabilization after using the medical product within the time frame of a clinical trial; and (3) not be expected to undergo change that is better explained by factors beyond the medical product. For example, concepts such as emotional and social impacts (e.g., feeling embarrassed or worried about cough; feeling that cough has annoyed friends or family) are distal and may be influenced by factors other than the condition or treatment, unlike proximal concepts such as symptoms and impacts on physical activities (e.g., chest or stomach pains; fatigue). In addition, it is best practice to evaluate each unique aspect of the identified concepts in isolation (e.g., symptom severity, symptom frequency, proximal impacts) to facilitate clear interpretation of the change in each concept as it relates to clinical benefit.

Whether a PRO is considered fit-for-purpose depends on the strength of the evidence (both qualitative and quantitative) demonstrating that the PRO score reflects the concept of interest within the target context of use. Sufficient evidence from existing literature and/or program-specific research must support the content validity and other quantitative psychometric properties of the selected PRO(s) for

[5] Fit-for-purpose: A conclusion that the level of validation associated with a tool is sufficient to support its context of use .
[6] Concept of interest: The concept is the aspect of an individual's clinical, biological, physical, or functional state, or experience that the PRO assessment is intended to capture or reflect .

the specific context of use. Qualitative data (e.g., from concept elicitation and cognitive debriefing interviews with patients) are essential for establishing the content validity of a PRO instrument. Content validity must be established before evaluating quantitative psychometric properties. Regulatory considerations on fit-for-purpose clinical outcome assessments (COAs), including PROs, are thoroughly described in FDA's draft guidance, Patient-Focused Drug Development: *Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments* (June 2022).[7]

An endpoint constructed from a fit-for-purpose PRO has the potential to support regulatory decision-making and labeling claims. Regulatory considerations on COA-based endpoints, including PRO-based endpoints, are thoroughly described in FDA's draft guidance, Patient-Focused Drug Development: *Incorporating Clinical Outcome Assessments Into Endpoints For Regulatory Decision Making* (April 2023).[8] The draft guidance highlights the importance of evaluating the meaningfulness of a treatment benefit to inform regulatory decision-making. Specifically, it is imperative that the change in the PRO-based endpoint reflects a change that is meaningful to patients. Anchor-based methods are the primary methods the FDA uses to interpret meaningful score changes in PRO-based endpoints. Qualitative methods such as cognitive interviews or exit interviews may be explored to complement the anchor-based methods in deriving a range of meaningful change thresholds.

6.4.2 Leicester Cough Questionnaire (LCQ)

One of the considerations when evaluating a PRO is that the change in the PRO score should be understandable and correspond to clinically meaningful improvement from the patient's perspective. For the LCQ Total Score, the Applicant proposed a responder thresholder of change of ≥1.3 points. The Applicant did not provide sufficient data to support the responder threshold of ≥1.3 points as meaningful to subjects. This cutoff was based on preliminary findings derived in a published study (Raj et al. 2009) with methodological concerns in the derivation of the cutoff that limit the utility of the 1.3-point change in the LCQ total score for use as a responder threshold, including:

- The use of the Global Rating of Change Questionnaire (GRCQ), a 15-point scale scored between +7 (a great deal better) and -7 (a great deal worse), to anchor a change in the LCQ score(s). The GRCQ is prone to recall bias and includes response options that are clinically indistinct and overlapping.
- It is unclear what constitutes a meaningful change on the GRCQ from the patient's perspective.
- The cutoff was derived as the change in the LCQ score corresponding to a small change in the GRCQ score, where small change was defined as scores of -3 ["Somewhat worse"], -2 ["A little worse"], 2 ["A little better"], 3 ["Somewhat better"] on the GRCQ. However, we have concerns about combining response categories indicating improvement and worsening to use as an anchor for deriving a meaningful change threshold, because thresholds for clinically meaningful improvement and worsening should be derived separately and are typically not symmetrical.

To inform the assessment of meaningful change in the LCQ total score for an individual patient, FDA requested that the Applicant explore responder analyses at additional thresholds. However, neither of the Applicant's proposed higher thresholds (≥3.3 or ≥4.1 points) have sufficient support to demonstrate that they reflect meaningful within-patient change. As a general rule, the measurement concept should be aligned between the anchor and the target PRO for the anchor-based analysis to be interpretable.

[7] When final, this guidance will represent the FDA's current thinking on this topic. For the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents
[8] When final, this guidance will represent the FDA's current thinking on this topic. For the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents.

Without evidence from the quantitative anchor-based analysis and/or qualitative evidence, clinically meaningful thresholds cannot be determined.

As stated previously, the FDA reviewed post hoc analyses on the LCQ physical domain because this domain directly assesses cough and related impacts. The Applicant's qualitative research in CC patients supported the content validity of the LCQ physical domain. While evidence from the published literature appears to show reasonable quantitative psychometric properties of the LCQ physical domain in CC patients (Birring et al. 2003; Nguyen et al. 2022), there are limitations with the literature findings (e.g., lack of details on the confirmatory factor analysis, low to moderate correlation with other conceptually related measures) that should be taken into consideration when applying the evidence to the current trials. Quantitative psychometric properties of the LCQ physical domain should be re-evaluated using additional data collected within the target context of use prior to concluding definitively that it is fit-for-purpose.

6.4.3 Cough Severity Diary Total Score

To support the content validity and psychometric properties of the CSD for use in this development program, the Applicant submitted evidence from the qualitative and quantitative development of the CSD for FDA review. The FDA generally agreed that the Applicant's developmental activities for the CSD were reasonable, in principle, to support its potential use as a secondary endpoint tool, but provided recommendations to the Applicant to further refine the development of the CSD.

For both trials, the Applicant prespecified a responder analysis of reduction from baseline in weekly CSD total score at the end of the main study period, selecting 2 different thresholds of ≥1.3 and ≥2.7-point reduction to define responders. During development, FDA requested that the Applicant provide evidence to demonstrate that the responder thresholds reflect meaningful change from the patient perspective; the Applicant has not addressed these concerns. Specifically, at the End of Phase 2 meeting, the FDA advised the Applicant to provide evidence and rationale that a 1.3-point reduction in CSD total score is clinically meaningful and appropriate to define responders. FDA questioned the use of "somewhat improved" on the PGIC as an anchor to define a clinically meaningful threshold of response (at ≥1.3 on the CSD). Those who rated themselves as "improved" or "very much improved" on the PGIC had a mean improvement of 2.7 points in the mean weekly CSD total score. The Applicant has not provided sufficient evidence to support each proposed responder threshold.

For completeness, the FDA considered post hoc item-level descriptive analyses and anchor-based analyses for the interpretation of meaningful change to contextualize the treatment effect. Item-level descriptive analyses (data not shown) showed a similar reduction in score in each CSD item for both the treatment and placebo arms, thus, indicating a small treatment effect. Anchor-based analyses were limited by the choice of a single anchor, the PGIC. Because each anchor has strengths and limitations (e.g., PGIC is prone to recall bias), the FDA recommends determining clinically meaningful improvement in PRO scores using multiple anchor scales, as was communicated to the Applicant during development. Using multiple anchors such as PGIC and PGIS would have allowed exploration of different aspects of the disease and characterization of the relationship among the various outcomes. As a single anchor, the PGIC's measurement concept (i.e., global change in cough) is not fully aligned with the measurement concept of the CSD (i.e., cough frequency, cough intensity, and disruption due to cough). Hence, data from these analyses are limited in their ability to establish whether the selected responder thresholders are clinically meaningful.

6.4.4 Cough Severity VAS

The VAS response scale has general limitations associated with its use and interpretation (i.e., possible inconsistencies with the length of the VAS due to formatting issues, relatively high rates of missing data or incomplete data with VAS, lack of clinically distinct response categories). For example, the Cough Severity VAS scale administered in the pivotal trials (Section 6.4.7) did not contain numbers on the scale; this omission introduces more uncertainty when trying to interpret the meaning of small numerical changes. As such, the FDA typically cautions sponsors against using the VAS response scale as an efficacy endpoint in clinical trials. At the End of Phase 2 meeting, the FDA advised the Applicant that we typically recommend using a numeric rating scale (NRS) (e.g., 0–10-point scale, anchored at both ends) or a simple verbal response scale (e.g., Likert-type scale), and stated that it was at the Applicant's discretion to use the Cough Severity VAS in the phase 3 trials or not.

Despite these concerns, the evaluated data from post hoc analyses of the VAS to characterize clinical benefit in the gefapixant trial. Descriptive analysis of the change in the weekly Cough Severity VAS score over time (data not shown) showed a similar change for both the treatment and placebo arms, thus, indicating a small treatment effect. In addition, data from post hoc anchor-based analyses for the interpretation of a meaningful change on the VAS is uncertain given the limitations with the use of a single anchor (PGIC), as described above in Section 6.4.3, due to the reasons described above for the Cough Severity Diary.

6.4.5 Copy of the LCQ

Figure 18. Copy of the LCQ (Adapted From Birring et al. (2003))

This questionnaire is designed to assess the impact of cough on various aspects of your life. Read each question carefully and answer by CIRCLING the response that best applies to you. Please answer ALL questions, as honestly as you can.

1. In the last 2 weeks, have you had chest or stomach pains as a result of your cough?

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

2. In the last 2 weeks, have you been bothered by sputum (phlegm) production when you cough?

1	2	3	4	5	6	7
Every time	Most times	Several times	Some times	Occasionally	Rarely	Never

3. In the last 2 weeks, have you been tired because of your cough?

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

4. In the last 2 weeks, have you felt in control of your cough?

1	2	3	4	5	6	7
None of the time	Hardly any of the time	A little of the time	Some of the time	A good bit of the time	Most of the time	All of the time

5. How often during the last 2 weeks have you felt embarrassed by your coughing?

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

6. In the last 2 weeks, my cough has made me feel anxious

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

7. In the last 2 weeks, my cough has interfered with my job, or other daily tasks

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

8. In the last 2 weeks, I felt that my cough interfered with the overall enjoyment of my life

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

9. In the last 2 weeks, exposure to paints or fumes has made me cough

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

10. In the last 2 weeks, has your cough disturbed your sleep?

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

11. In the last 2 weeks, how many times a day have you had coughing bouts?

1	2	3	4	5	6	7
All of the time (continuously)	Most times during the day	Several times during the day	Some times during the day	Occasionally through the day	Rarely	None

12. In the last 2 weeks, my cough has made me feel frustrated

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

13. In the last 2 weeks, my cough has made me feel fed up

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

14. In the last 2 weeks, have you suffered from a hoarse voice as a result of your cough?

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

15. In the last 2 weeks, have you had a lot of energy?

1	2	3	4	5	6	7
None of the time	Hardly any of the time	A little of the time	Some of the time	A good bit of the time	Most of the time	All of the time

16. In the last 2 weeks, have you worried that your cough may indicate serious illness?

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

17. In the last 2 weeks, have you been concerned that other people think something is wrong with you, because of your cough?

1	2	3	4	5	6	7
All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

18. In the last 2 weeks, my cough has interrupted conversation or telephone calls

1	2	3	4	5	6	7
Every time	Most times	A good bit of the time	Some of the time	A little of the time	Hardly any of the time	None of the time

19. In the last 2 weeks, I feel that my cough has annoyed my partner, family or friends

1	2	3	4	5	6	7
Every time I cough	Most times when I cough	Several times when I cough	Some times when I cough	Occasionally when I cough	Rarely	Never

Thank you for completing this questionnaire.

Source: Birring et al. (2003).

Abbreviation: LCQ, Leicester Cough Questionnaire

6.4.5.1 Conceptual Framework of the LCQ

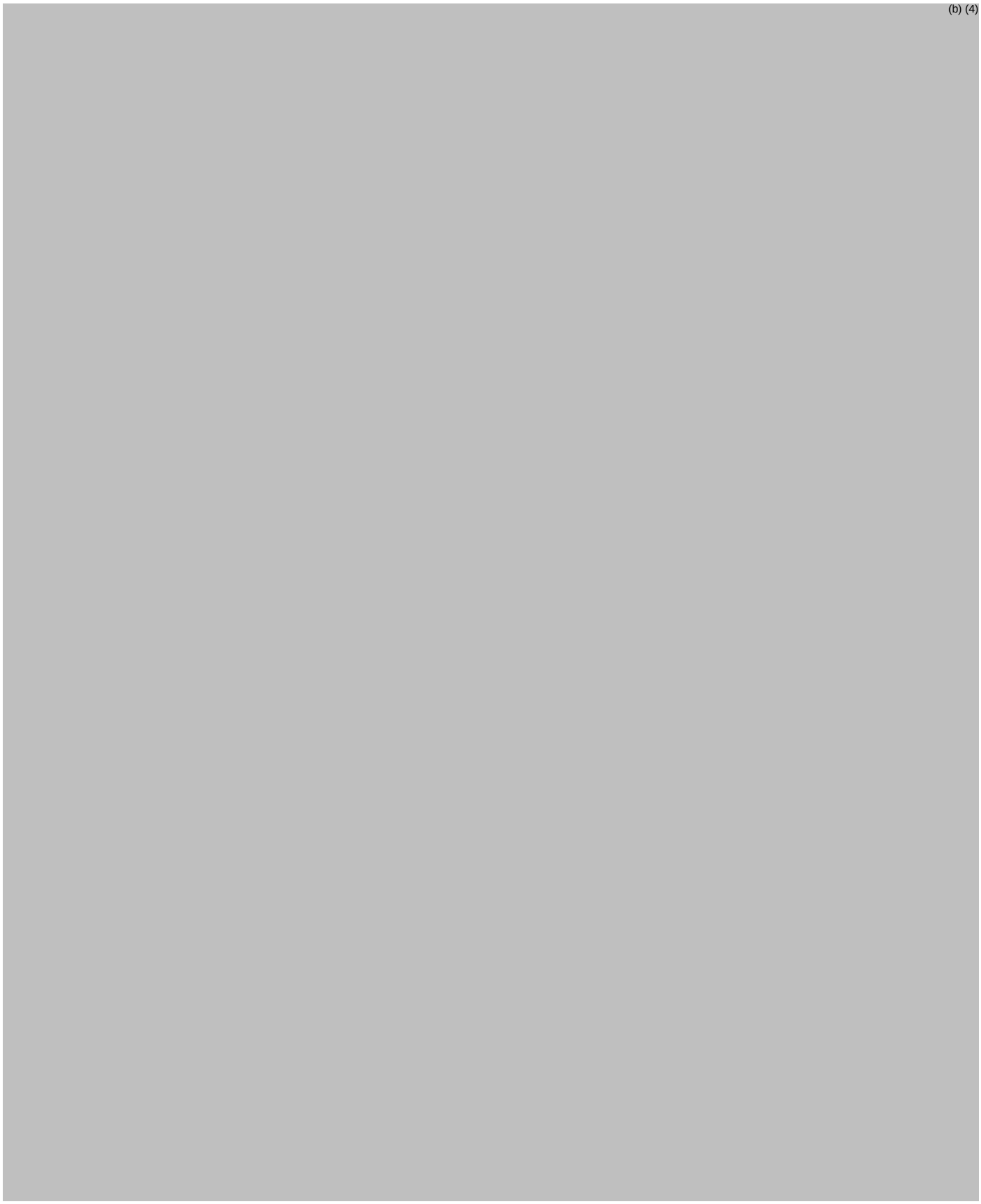
Table 44. Conceptual Framework of the LCQ

Domain	Items	General Concept
Physical	**Item 1:** In the last 2 weeks, have you had chest or stomach pains as a result of your cough?	Cough-related symptoms and their physical impacts
	Item 2: In the last 2 weeks, have you been bothered by sputum (phlegm) production when you cough?	
	Item 3: In the last 2 weeks, have you been tired because of your cough?	
	Item 9: In the last 2 weeks, exposure to paints or fumes has made me cough	
	Item 10: In the last 2 weeks, has your cough disturbed your sleep?	
	Item 11: In the last 2 weeks, how many times a day have you had coughing bouts?	
	Item 14: In the last 2 weeks, have you suffered from a hoarse voice as a result of your cough?	
	Item 15: In the last 2 weeks, have you had a lot of energy?	
Psychological	**Item 4:** In the last 2 weeks, have you felt in control of your cough?	Cough-related psychological impacts
	Item 5: How often during the last 2 weeks have you felt embarrassed by your coughing?	
	Item 6: In the last 2 weeks, my cough has made me feel anxious	
	Item 12: In the last 2 weeks, my cough has made me feel frustrated	
	Item 13: In the last 2 weeks, my cough has made me feel fed up	
	Item 16: In the last 2 weeks, have you worried that your cough may indicate serious illness?	
	Item 17: In the last 2 weeks, have you been concerned that other people think something is wrong with you, because of your cough?	
Social	**Item 7:** In the last 2 weeks, my cough has interfered with my job, or other daily tasks	Cough-related social impacts
	Item 8: In the last 2 weeks, I felt that my cough interfered with the overall enjoyment of my life	
	Item 18: In the last 2 weeks, my cough has interrupted conversation or telephone calls	
	Item 19: In the last 2 weeks, I feel that my cough has annoyed my partner, family or friends	

Source: Reviewer.

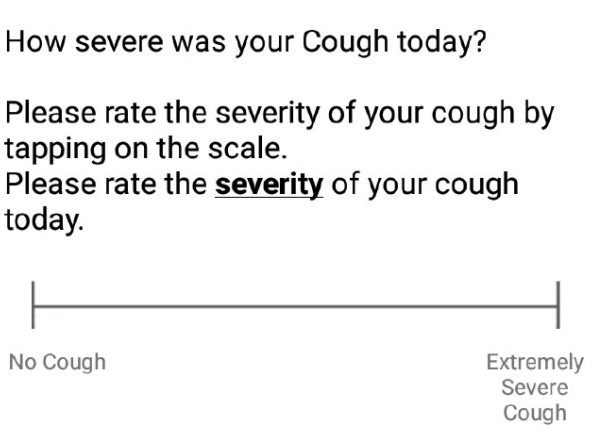
6.4.6 Copy of the CSD

(b) (4)



6.4.7 Copy of the Cough Severity VAS

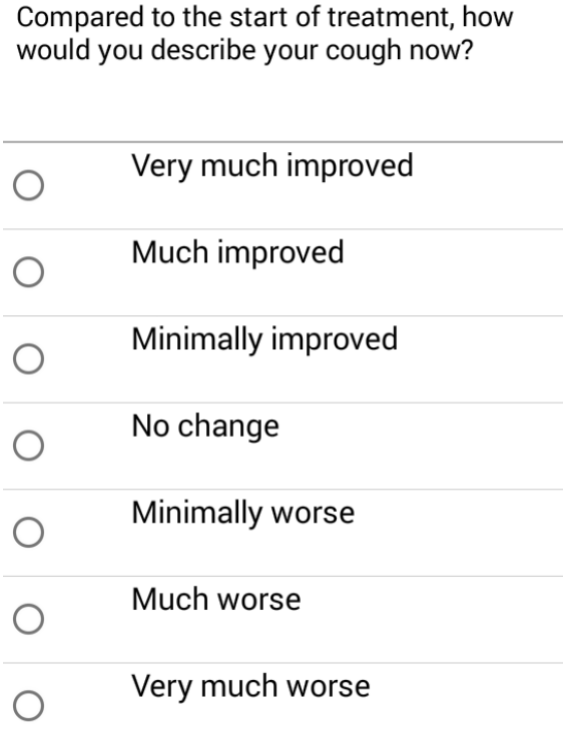
Figure 20. Copy of the Cough Severity VAS



Source: Appendix 5 and Appendix 6 of the Applicant's response to the August 9, 2023, Information Request (1 of 2).
Abbreviation: VAS, visual analog scale

6.4.8 Copy of the PGIC in Trials P027 and P030

Figure 21. Copy of the PGIC in Trials P027 and P030



Source: Appendix 1 of the Applicant's response to the FDA Information Request dated February 26, 2021.
Abbreviations: FDA, Food and Drug Administration; PGIC, patient global impression of change