U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	sNDA 207986
Supplement #:	Supplement-2
Drug Name:	OTIPRIO (ciprofloxacin otic suspension)
Indication(s):	Acute Otitis Externa
Applicant:	Otonomy, Inc.
Date(s):	Receipt date: May 2, 2017 Due date: March 2, 2018
Review Priority:	Standard
Biometrics Division:	OB DBIV
Statistical Reviewer:	Ed Bein
Concurring Reviewers:	Karen Higgins
Medical Division:	DAIP
Clinical Team:	Mark Needles
Project Manager:	Jane Dean

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

This review evaluates the strength of the evidence for the efficacy and safety of OTIPRIO for the treatment of acute otitis externa (AOE), based on analyses of data from a Phase 3 multisite double-blinded randomized controlled trial involving 262 participants. It concludes that there is very strong evidence for the efficacy of OTIPRIO and finds no evidence of safety concerns.[1] On this basis, it recommends that the drug be approved for the indication of AOE.

The analysis of efficacy was based on data from the 260 participants (130 in the OTIPRIO arm, 130 in the control arm) in the intention-to-treat (ITT) population. This population included all participants who were randomized and did not have Group A Streptococci cultured on the first day of the trial. Participants in the OTIPRIO arm received 12 mg of the drug in the affected ear on day 1, while participants in the control arm received a sham (air in syringe) administration of the drug. Participants were followed for a month. The primary efficacy endpoint, clinical cure, was assessed at day 8 (study visit 3). Clinical cure was defined by the sponsor as the absence of edema, erythema, and tenderness. Secondary efficacy endpoints included clinical cure at days 4 (study visit 2) and 15 (study visit 4). In addition, other secondary endpoints involved clinical cure at different time points with respect to a subset of the ITT population. The microbiological intention-to-treat (Mic-ITT) population included ITT participants who had a positive culture for *S. aureus* or *P. aeruginosa* at study entry. This population included 52 participants in the OTIPRIO arm and 56 in the control arm.

In addition to the sponsor's definition of the primary efficacy endpoint, several alternative definitions were also analyzed. These alternative definitions embodied different constellations of signs and symptoms that are commonly found in AOE patients. One alternative definition of clinical cure that was examined is the absence of edema, erythema, tenderness, and otorrhea. This adds the sign of otorrhea to the sponsor's definition, and is the preferred definition of the FDA clinical review team. All ITT analyses of the sponsor's definition, of the preferred definition, and of the other alternative definitions of the primary endpoint found statistically significant superiority of the use of OTIPRIO vs. no treatment (see Table 4). In addition, statistically significant superiority of OTIPRIO vs. no treatment was also found for clinical cure (both the sponsor's definition and the preferred definition) at day 15 on the ITT population and days 8 and 15 on the Mic-ITT population (see Tables 5 and 6). In sum, analyses of the primary and secondary efficacy endpoints provide very strong evidence of the superiority of OTIPRIO vs. no treatment at days 8 and 15.

Analyses of safety data involved the 259 participants (127 in the OTIPRIO arm, 132 in the control arm) who received study drug (see Table 7). 54 participants experienced at least one adverse event, almost two-thirds of which were mild. Descriptive statistics comparing adverse event incidence in the two arms yielded no evidence of safety concerns regarding the use of OTIPRIO.

[1] Please refer to the clinical reviewer's review for a more in-depth analysis of safety data.

Given the results of the analyses of the efficacy and safety data, approval of OTIPRIO for the treatment of AOE is recommended.

2 INTRODUCTION

2.1 Overview

OTIPRIO is a 6% ciprofloxacin otic suspension. It was approved by the FDA for the treatment of pediatric patients with bilateral otitis media with effusion undergoing tympanostomy tube placement (FDA approval letter December 10, 2015). In IND110244, its sponsor subsequently proposed the additional use of OTIPRIO for the treatment of acute otitis externa (AOE) in children and adults. This was based on the fact that the antibiotic ciprofloxacin is known to be effective in treating the strains *S. aureus* or *P. aeruginosa*, which are often associated with AOE. In a letter of April 14, 2016, the FDA agreed with the additional indication of AOE and further agreed with the sponsor that a single Phase 3 multicenter randomized double-blind sham-controlled trial would suffice to demonstrate the efficacy and safety of OTIPRIO for the indication of AOE.

The sponsor's Phase 3 trial, named Trial 201-201609, is the subject of this review. Table 1 provides a summary of the study.

Table 1: List of studies included in analysis

Applicant-defined study number	Phase and Design	Treatment Period	Follow-up Period	# of Subjects per Arm	Study Population
201-201609	Phase 3	Single dose, administered on study day 1	29 days	130[1]	Children, adolescents, and adults with AOE & without positive baseline culture for group A streptococci at study entry.

[1] Two participants who were randomized proved to have positive baseline cultures for group A streptococci at study entry and were excluded from efficacy analyses. This resulted in 130 participants per arm who were included in efficacy analyses.

2.2 Data Sources

This review is based on material presented in the Phase 3 trial's March 31, 2017 Clinical Study Report (CSR), the trial's December 12, 2016 final Statistical Analysis Plan (SAP), and on data contained in ADAM data sets. The CSR is given at \\CDSESUB1\evsprod\NDA207986\0077\m5\53-clin-stud-rep\535-rep-effic-safety-stud\acute-otitis-externa\5351-stud-rep-contr\201-201609\201-201609-body.pdf. The SAP is given at \\CDSESUB1\evsprod\NDA207986\0077\m5\53-clin-stud-rep\535-rep-effic-safety-stud\acute-

otitis-externa\5351-stud-rep-contr\201-201609\201-201609-e3-16-1-09.pdf. The data sets are contained in folder \\CDSESUB1\evsprod\NDA207986\0077\m5\datasets\201-201609\analysis\adam\datasets. The data sets analyzed are adyo.xpt, adsl.xpt, adyc.xpt, adcm.xpt, adyd.xpt, and adae.xpt.

3 STATISTICAL EVALUATION

This section presents a detailed review of the statistical analyses of primary and secondary efficacy endpoints and safety data from the Phase 3 trial 201-201609. Reviewer's comments on the adequacy of the study design and sponsor's analysis are given in italics.

3.1 Data and Analysis Quality

The sponsor's ADAM data sets were adequately documented, and it was possible to straightforwardly replicate the sponsor's data analyses of the primary and secondary efficacy endpoints, as presented in the SAP and CSR. The SAP was finalized prior to unblinding. In the course of replicating these analyses, the reviewer discovered that the sponsor had incorrectly categorized one participant on the primary efficacy endpoint (confirmed by the sponsor: see Table 4, note 1), and analyses of the primary endpoint used the correct categorization.

As noted above, the reviewer created several alternative versions of the primary efficacy endpoint, based on alternative constellations of signs and symptoms associated with AOE. The data analyses of these alternative endpoints used the same statistical methods as used to analyze the sponsor's version of the endpoint (per the SAP). Similarly, multiple versions of some of the secondary efficacy endpoints were created, and again the statistical methods used were the same as those described in the SAP.

The sponsor also specified a time-to-cessation-of-otalgia secondary efficacy endpoint, and no attempt was made to replicate the sponsor's analysis here. As discussed below, the reviewer believes that these analyses are flawed, and instead performed alternative time-to-cessation analyses.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

The Phase 3 trial (Trial 201-201609) was a multisite two-arm double-blind sham-controlled randomized trial. Eligible participants at 37 sites in the United States were randomized (1:1 ratio, using randomly permuted blocks) to either receive a single 12 mg dose of OTIPRIO or single sham (empty syringe) dose in the study ear on day 1, administered by an unblinded clinician, and then followed for four weeks.[2] Inclusion criteria included age 6 months or older with a clinical

[2] If a participant was affected in both ears, then he/she was supposed to receive the assigned treatment in both ears. There was a small number of such cases in which only the study ear was treated.

diagnosis of AOE in at least one ear. Exclusion criteria included tympanic membrane perforation, chronic OE, and use of antimicrobial ear drops within one week of the screening visit. Patients with a positive baseline culture for group A streptococci were removed from the efficacy evaluation, administered standard of care for AOE, but included in the safety evaluation. For patients with both ears affected, the SAP includes an algorithm for determining the study ear; the more severely affected ear, unless both are equally affected, in which case the right ear is the study ear.

At study visits on days 1, 4, 8, 15, and 29 (corresponding to visits 1, 2, 3, 4, and 5, respectively), a blinded clinician performed an otoscopic exam to evaluate the severity of edema, erythema, and tenderness on a 4-point scale (0=absent, 3=severe), upon which the primary efficacy endpoint was based. At this exam, the clinician also evaluated the presence and severity of otorrhea. At visit 1 and at all subsequent visits at which at least one of otorrhea, edema, erythema, or tenderness was noted, a culture was taken from the study ear. In addition, participants aged 3 years and older (or their caregivers) recorded otalgia level in a daily diary using the Wong-Baker FACES Pain Rating Scale (0=no hurt, 10=hurts worst) during days 1 through 15. Any participant who failed to show improvement on the otoscopic exam by visit 3 or whose daily pain diary showed no improvement was provided the standard of care for AOE at the discretion of the site's investigator.

The intention-to-treat (ITT) sample included all randomized participants who did not have a positive baseline culture for group A streptococci. Two of the 262 participants who were randomized were excluded from the ITT sample due to such a positive baseline culture. Hence, the ITT sample included 260 participants, evenly divided between the two arms. The number of ITT participants per site ranged from 1 to 31; median = 3. The microbiological intention-to-treat (Mic-ITT) sample was the subsample of ITT participants with a positive baseline culture for *S. aureus* or *P. aeruginosa*. The Mic-ITT sample included 108 participants, with 52 assigned to the active treatment arm and 56 assigned to the sham arm.

Clinical cure (CC) at a given visit was defined as the absence of edema, erythema, and tenderness in the study ear, without the concomitant use of other systemic or local (to the study ear) antibiotics. The day of the visit must fall within a time window specified in the SAP; visits falling outside of the window are not considered an official study visit and any data collected during the visit are irrelevant to defining CC. Concomitant use means any use after the administration of the study drug and prior to or on the day of the visit. The primary efficacy endpoint was CC at day 8 using the ITT sample.

The six secondary efficacy endpoints were (1) CC at day 15 using the ITT sample, (2) CC at day 8 using the Mic-ITT sample, (3) CC at day 15 using the Mic-ITT sample, (4) CC at day 4 using the Mic-ITT sample, (5) CC at day 4 using the ITT sample, and (6) time to cessation of otalgia, without using analgesics to treat the study ear at any time through day 17[3] (TTCO) using the subset of ITT participants who completed the daily diary at at least one study visit and who satisfy neither of the following conditions

[3] The CSR and SAP sometimes refer to TTCO as time to cessation through 15 days and sometimes through 17 days.

- No use of concomitant analgesics and no otalgia through day 17.
- No use of concomitant analgesics and there is a day with a missing pain rating that is in between two days with a rating of no pain.

Reviewer Comments

There are grounds for questioning the adequacy of the definition of clinical cure; several of these grounds were communicated to the sponsor prior to the start of the Phase 3 trial[4]. As it is defined by the sponsor,

1. *CC takes into account prior use of antibiotics but not of analgesics. However, analgesics can reduce tenderness, and the clinician's tenderness rating is one determinant of CC. Hence, for participants who used analgesics, being rated a CC treatment success at a given time point could reflect the effects of both the study medication and concomitant analgesics.*
2. *CC does not take into account the presence of otorrhea, which is rated by the clinician, nor of otalgia, which is rated by the participant during the first 15 days. This means that a participant can be rated a CC treatment success at a given time point while exhibiting otorrhea and/or experiencing otalgia. This is, then, a circumscribed notion of clinical cure.*
3. *CC treatment success for a participant at a given day was sometimes followed by CC treatment failure at a succeeding day. It is possible that, for some participants, this reflects the initial AOE bout being completely cured but then the participant acquiring a new case of AOE. However, given the limited time span participants were followed, it seems more likely that CC treatment success followed by treatment failure instead indicates that the rating of treatment success did not represent a durable AOE cure.*

As discussed in the Results section below, the reviewer conducted sensitivity analyses to examine more robust definitions of clinical cure.

There are also grounds for questioning the adequacy of the definition of TTCO and of the sample used for survival curve estimation and testing. As defined by the sponsor, TTCO takes into account use of analgesics but not of antibiotics. However, concomitant antibiotics can reduce otalgia. Hence, for participants who used concomitant antibiotics, having a TTCO value less than or equal to 15 days could reflect the effects of both the study medication and concomitant antibiotics. Instead, TTCO should reflect the time till cessation while adherent to the treatment protocol with respect to concomitant antibiotics use. Finally, the sample employed leads to "apples-and-oranges" comparisons. Consider, for example, a participant who would experience no otalgia if assigned to the active treatment group but who would experience otalgia on day 3 if assigned to the sham group. If this participant were in fact assigned to the active treatment, then he/she would be excluded from the sample used for survival curve estimation and testing, while if

[4] An August 19, 2016 letter from the FDA project manager to the sponsor stated, "Please clarify why the signs and symptoms of OE scoring scale will not be used to assess the sign "otorrhea" or the symptom "otalgia" as done in the preceding Phase 2 study. We recommend you consider incorporating the assessment of otorrhea into the primary efficacy endpoint."

the participant were in fact assigned to the sham treatment, then he/she would be included in the sample. Treatment effects cannot be estimated without bias given such an "apples and oranges" sample. Also, the excluding conditions used to specify the sample for TTCO analysis are ambiguous, as it is unclear whether a participant rating the first k<15 days as without pain but followed by the remaining days missing pain ratings is to be excluded. An alternative TTCO analysis that avoids these difficulties is presented below.

3.2.2 Statistical Methodologies

Estimation and Testing

With the exception of time to cessation of otalgia, the primary and secondary efficacy endpoints are binary variables. The parameter of interest to be estimated and tested is the population difference in the proportions of successes between the active treatment and sham conditions. For each binary endpoint, a point estimate for the difference in proportions was computed and an unconditional 95% confidence interval (CI) for two independent binomial observations was constructed for the difference in proportions.[5] The null hypothesis that the population difference in proportions equals zero was tested with a 2-sided Fisher exact test.[6] In addition, the population proportion of successes was estimated for each arm and the Clopper-Pearson method was used to construct 95% CIs.

For TTCO, up to day 15, each arm's survival curve was estimated by the Kaplan-Meier estimator. The log-rank test was used to test the null hypothesis that at all times up to day 15, the hazard rates for the two arms are identical.

No baseline covariates were used in any hypothesis testing.

Controlling the Familywise Error Rate (FWER)

A gatekeeping strategy was followed. This is a closed testing procedure that maintains an FWER set at .05. Contingent on the test of the primary efficacy endpoint using the ITT sample yielding a statistically significant p-value at 2-sided alpha = .05, the six secondary endpoints were tested in the order they are listed above, using the appropriate sample, with 2-sided alpha = .05. If any of these tests yielded a nonsignificant p-value, then the succeeding endpoint test results were reclassified as exploratory.

---

[5] Technical note: The sponsor computed unconditional CIs using the *exact riskdiff* statement in SAS proc freq. The reviewer computed unconditional CIs using the function *uncondExact2x2* (with arguments *method = "simple", tsmethod="central"*) from the R package *exact2x2*. Both software gave very similar results, with the R function's CIs being contained within the SAS CIs. For a general discussion of unconditional CIs, see Agresti, A. (2013). *Categorical Data Analysis*. 3rd ed., page 609.

[6] Technical note: the Fisher exact test is a conditional test, but the 95% CIs were unconditional. This implies that it is possible that hypothesis testing could yield a significant p-value while, at the same time, the 95% CI included zero, or, conversely, the unconditional 95% CI could exclude zero while the Fisher exact test p-value is not significant.

Handling Missing Data

There were no missing CC data at day 1. Missing CC endpoint data for days 4, 8, 15, and 29 were singly imputed. If a participant withdrew from the study without the given day's clinical exam, then he/she was considered a treatment failure for that day's visit. Otherwise,

- For days 8 and 15, the CC value imputed (i.e., treatment success or treatment failure) was determined by the CC status of the immediately preceding and immediately succeeding study visit. "Success" was imputed if the CC status at preceding and succeeding visits is "Success;" else "Failure" was imputed.
- For day 4, "Failure" was imputed.
- No imputation was made for day 29.

The SAP does not explain how to impute values for CC at days 8 or 15 when the participant was not a dropout and either the preceding or succeeding study visit was not attended.

For the participants included in the TTCO analyses, TTCO was censored at 17 days if either of the following occurred

- There was no cessation of otalgia.
- An analgesic was used for the treatment of pain in the study ear.

If a participant has cessation of otalgia by day $k<15$ followed by missing pain ratings for the remaining days, then TTCO = k.

The SAP does not explain, for example, how to handle the case of no analgesic use, pain ratings for the first k days, and missing pain ratings for the remaining days. There are other missing data scenarios that are not covered as well.

Missing otorrhea data were handled analogously to missing CC data.

Handling missing microbiological data was not addressed in the SAP.

Sensitivity Analyses

The TTCO was recalculated by ignoring any concomitant analgesic use and Kaplan-Meier curves were computed.

Reviewer Comments

There are limited CC missing data, as documented below, and therefore even inadequate methods of handling such missing data may not have dramatically negative consequences for estimation and testing. Nonetheless, the sponsor's proposed handling of missing data is limited. It uses single imputation, which often leads to underestimated standard errors and p-values. Its use of single imputation is especially concerning because its deterministic algorithm for imputing values may generate extremely biased estimates of arm-specific treatment success

rates. The algorithm was evaluated by examining the observed CC values at day 8 of the 191 ITT participants who hadn't previously used concomitant antibiotics and who were not study dropouts. Under the pretense that these day 8 observed CC values were not observed, the algorithm was used to impute values for the 191 participants, and then the imputed values were compared to the observed values. In four cases, the algorithm failed to generate an imputed value. With the remaining 187 cases, arm-specific success proportions and the difference in proportions were computed twice, once using the observed values and once using the imputed values. Using the 187 observed values, the success proportions of the active treatment and sham arms were 83.5% and 73.8%, respectively, giving a difference in proportions of 9.7%. Using the corresponding imputed values, the analogous proportions were 53.4% and 44.0%, giving a difference in proportions of 9.4%. Hence, using either the observed or imputed values yielded about the same difference in proportions for this subsample of 187 participants, but use of the imputed values yielded negative biases of about 30% for the arm-specific success proportions. While it is reassuring that bias was not introduced for the difference in proportions, it is disquieting that use of the algorithm introduced such substantial bias for the arm-specific success proportions. It is not possible to know the magnitude of biases, if any, that might be introduced when the algorithm is applied to participants with actually unobserved day 8 CC values, and this would also depend on these participants' mix of day 4 and day 15 CC values and on the extent of missingness in each arm. On the face of it, however, it is not clear the algorithm is dependable.

As noted, the sponsor did not perform any missing data sensitivity analyses. Alternatively, data analyses could be performed with missing values filled in under a "worst case" scenario (WCS), as this would give a "worst case" bound on estimates of and test results for the treatment effect and arm-specific success proportions. In a WCS, missing endpoint values for active treatment participants are filled in as treatment failures, while missing endpoint values for sham participants are filled in as treatment successes. This results in the smallest possible estimate of the difference in proportion of successes that is consistent with the observed data. Then, in testing the difference in proportions, the p-value obtained from a WCS analysis would be an upper bound for the p-value that would be obtained from an (infeasible) analysis that used the true values of the missing observations. Hence, if a WCS analysis yielded a statistically significant p-value, then the infeasible analysis using the true endpoint values for all participants would necessarily also yield a significant p-value. The reviewer gives the results of WCS analyses below.

There are greater concerns about the adequacy of the TTCO data analyses. First, it is questionable to set TTCO = k when otalgia ceases by day k<15 followed by all missing pain ratings. Second, because analgesics are very short-acting and are palliative rather than curative, the use of analgesics should not necessarily lead to a value of 17 being assigned to TTCO, as the sponsor does. For example, suppose a participant used analgesics through study day 3 but not thereafter, and from day 5 on records diary pain ratings of 0. The use of analgesics up to day 3 does not impact pain level at day 5 or subsequently, and therefore it is appropriate to consider pain having ceased by day 5 (assuming no use of concomitant antibiotics). Third, the standard Kaplan-Meier estimation and log rank tests the sponsor used give unbiased results under the assumption that missing pain ratings from the 15-day diaries are missing at random (MAR). This assumption would not hold, for example, if days during which no pain was

experienced were more likely to go unrated than days during which pain was experienced (or vice versa). The sponsor does not address whether the MAR assumption is plausible, nor perform sensitivity analyses examining robust of results to violations of the MAR assumption.

3.2.3 Patient Disposition, Demographic, and Baseline Characteristics

Table 2 compares the active treatment (130 participants) and sham (130 participants) arms from the ITT sample on baseline variables, giving standardized differences between the arms. A standardized difference is the difference in means between the arms divided by a pooled standard deviation term. Due to the use of random assignment, in large samples these differences should approach zero, but in more modest samples non-zero differences are to be expected. On demographics variables, the largest standardized difference was on the proportion of participants in the arm who were age 18 or older (77.7% for active treatment, 70.8% for sham).[7] On baseline health status variables, the largest standardized difference was on edema severity (on a 0-3 scale, mean of 1.78 for active treatment, mean of 1.71 for sham).

[7] For a binary variable, the standardized difference for the proportions with a value of 1 equals the standardized difference for the proportions with a value of 0. Hence, the standardized difference for the proportions of non-adults equals the standardized difference for the proportions of adults.

Table 2: Comparing Active Treatment and Sham Arms on Baseline Characteristics

Variable	Active Treatment Arm	Sham Treatment Arm	Standardized Difference[1]
gender			
female	57.7%	57.7%	0.0
age			
age in years	36.7	34.8	.10
adult (18+)	77.7%	70.8%	.16
race and ethnicity[2]			
Caucasian	55.4%	51.5%	.08
Afro-American	10.8%	11.5%	-.02
Latino	31.5%	33.1%	-.03
signs and symptoms[3]			
edema	1.78	1.71	.12
erythema	1.92	1.95	-.05
tenderness	1.99	1.96	.05
otorrhea	60.0%	58.5%	.03
bacterial infection[4]			
in Mic-ITT	40.0%	43.1%	-.06

Notes. All variables are assessed at baseline. The ITT sample include 130 participants in each arm.
[1] The standardized difference is the difference between the means in the two arms (for a binary variable, the difference in proportions) divided by the square root of a pooled standard deviation term. It gives the effect size difference between the two arms.
[2] "Caucasian" denotes non-Latino and white. "Afro-American" denotes non-Latino and black.
[3] Edema, erythema, and tenderness were assessed on a 0-3 scale, with 0 indicating absence of the sign/symptom. These three ratings were used to determine the sponsor's clinical cure efficacy endpoint. All participants had at least one non-zero rating at baseline. The presence of otorrhea is represented by a binary variable.
[4] Membership in the Mic-ITT sample required a positive baseline culture for *S. aureus* or *P. aeruginosa*.

Seven participants dropped out of the study at some point after visit 1, and all seven were in the active treatment arm. Table 3 presents the amount of missing clinical cure endpoint data in the ITT sample for each of the five study visits. There are no missing endpoints for visit 1, and the active treatment arm has more missing endpoint data at all other visits. The missing data are due to study dropout or to intermittent missing of scheduled study visits, where a scheduled visit is considered missed if it did not occur within the time window specified for the visit in the SAP. Note that a missed visit does not necessarily cause missing clinical cure endpoints, since if concomitant antibiotics are used during the study, then the participant is considered a treatment failure at all subsequent study visits, whether or not they are attended.

Table 3: Missing Clinical Cure Endpoint Data in ITT Sample at Different Study Visits

Time Point	Active Treatment	Sham Treatment
Study Visit 1/Day 1	0	0
Study Visit 2/Day 4	8	1
Study Visit 3/Day 8	4	0
Study Visit 4/Day 15	9	3
Study Visit 5/Day 28	7	7

Notes. 130 participants in each treatment arm.

3.2.4 Results and Conclusions

Results for the primary and secondary efficacy endpoints are presented below.

Primary Efficacy Endpoint

Table 4 presents results of analyses of the sponsor's CC endpoint at day 8 on the ITT sample, along with results of analyses of seven modifications of this endpoint, which are referred to as sensitivity endpoints. These latter analyses are sensitivity analyses that explore (i) a different approach to handling missing data, and/or (ii) different plausible definitions of what it means to achieve clinical cure of AOE. The different versions of the CC endpoint that are examined in the sensitivity analyses are

1. CC_{ot}: sponsor's CC plus presence of otorrhea: a participant exhibiting otorrhea on day 8 is considered a treatment failure; in the absence of otorrhea, a participant's day 8 endpoint value is the same as that given by the sponsor's CC.
2. CC_a: "Worst case scenario" (WCS). Missing values in the day 8 CC endpoint are replaced by "treatment success" if the participant belongs to the sham arm and are replaced by "treatment failure" if the participant belongs to the active treatment arm. If the active treatment is statistically significantly superior to sham using the WCS version, then day 8 CC with missing values replaced by their true but unknown values would necessarily also show statistically significant superiority.
3. CC_b: CC_a plus whether otorrhea was present on day 8. This differs from CC_{ot} in that missing endpoint values are filled in via WCS rather than using the algorithm in the sponsor's SAP.
4. CC_c: CC_b plus whether analgesics were used on day 8. That is, participants who exhibited otorrhea or used analgesics on day 8 are considered treatment failures; otherwise, the endpoint value from CC_b is carried over. WCS is used to handle missing data.
5. CC_d: CC_a plus whether analgesics were used on day 8. That is, participants who used analgesics on day 8 are considered treatment failures; otherwise, the endpoint value from CC_a is carried over. WCS is used to handle missing data.
6. CC_e: CC_d plus whether relapse by day 15 (visit 4): first, the day 15 analogue to CC_d is computed. Second, "treatment success" is changed to "treatment failure" for day 8 CC_d if the day 15 CC_d analogue is "treatment failure." This is done because, if there was a

relapse by day 15, then the apparent treatment success at day 8 was very likely not a durable clinical cure. In sum, CC_e considers analgesic use and relapse by day 15, using WCS to handle missing data.

7. CC_f: CC_e plus whether otorrhea or otalgia were present on day 8 or otorrhea was present at day 15 (visit 4): the presence of otorrhea at day 15 represents a relapse; otalgia at day 15 was not examined because many participants had stopped rating daily otalgia by the time visit 4 occurred. CC_e "treatment success" was changed to "treatment failure" if this kind of relapse occurred or if otorrhea or otalgia were present on the day of visit 3.

Of the seven sensitivity endpoints, the FDA clinical review team considers CC_{ot} to be the most clinically relevant; indeed, the results for CC_{ot} are reported in the label (presented later).

There were 4 participants who had missing data for study visit 3, all in the active treatment arm. Both the SAP algorithm for imputing missing data for CC and the sensitivity analysis WCS algorithm for CC_a yielded the same endpoint values, so a separate row for CC_a is not included in Table 4 below. Similarly, CC_{ot} and CC_b had the same endpoint values, so a separate row for the latter is not included in the table.

Table 4: Results for Primary Efficacy Endpoint (Clinical Cure at Day 8/Visit 3) and Related Sensitivity Analyses for the ITT Sample

Version of Clinical Cure at Study Day 8 (Study Visit 3)	**Proportion of Successes in Active Tx**	**Proportion of Successes in Sham Tx**	**Difference in Proportions**
Sponsor's Version: No Edema/Erythema/ Tenderness w/o Antibiotics (CC)[1,2]	70.8% (62.2, 78.4)	48.5% (39.6, 57.4)	22.3% (10.4, 34.2) p=.0004
Sponsor's Version + No Otorrhea (CC_{ot})[3]	69.2% (60.5, 77.0)	46.2% (37.4, 55.1)	23.1% (10.7, 34.6) p=.0003
Sponsor's Version + No Otorrhea + No Analgesics (CC_c)	68.5% (59.7, 76.3)	46.2% (37.4, 55.1)	22.3% (10.1, 34.1) p=.0004
Sponsor's Version + No Analgesics (CC_d)	70.0% (61.3, 77.7)	48.5% (39.6, 57.4)	21.5% (9.1, 33.1) p=.0006
Sponsor's Version + No Analgesics + No Day 15 Relapse (CC_e)	63.8% (55.0, 72.1)	46.9% (38.1, 55.9)	16.9% (4.4, 28.7) p=.0087
Sponsor's Version + No Analgesics + No Day 15 Relapse + No Otorrhea + No Otalgia (CC_f)	45.4% (36.6, 54.3)	25.4% (18.2, 33.8)	20.0% (7.9, 32.0) p=.0011

Notes. The active treatment arm contained 130 participants and the sham arm also contained 130 participants. 4 participants in the active treatment arm and 0 in the sham arm had missing visit 3 data. Each cell in the table contains a point estimate and a 95% confidence interval (CI). CIs for proportion of successes are computed via the Clopper-Pearson method, and for the difference in proportions unconditional CIs for two independent binomial observations are constructed using the R statistical software (see footnote 7). The cells for difference in proportions also include the 2-sided p-value from a Fisher's exact test, using alpha = .05. CC_{ot}-CC_f are defined in the text above.

[1] The reviewer discovered that the sponsor assigned an incorrect CC value to one participant (Subject ID 200-9002) in the active treatment arm. The CC and sensitivity results reported in Table 4 use the corrected value for the participant, which changes the participant from a treatment failure to a treatment success.

[2] CC and the sensitivity endpoint CC_a took the same values and therefore a row for CC_a is not included in the table.

[3] CC_{ot} and CC_b took the same values and therefore a row for CC_b is not included in the table.

For CC at day 8 and for all seven sensitivity endpoints, the active treatment is superior to sham at $p<.01$ over the ITT sample. The results for CC through CC_d are very similar. Regarding CC_{ot}, 2 active treatment participants and 3 sham participants were assessed as treatment failures solely due to the presence of otorrhea; that is, these five participants had no tenderness, edema, or erythema, just otorrhea. CC_e, which additionally takes into account relapse at visit 4, shows a noticeable dip in the active treatment arm's success rate, with a consequent decrease in the estimated difference in proportions and increase in the p-value. There were 10 relapses at visit 4, 5 observed and 5 due to WCS filling in missing values for active treatment participants.

Regarding CC_f, additionally examining otalgia at visit 3 or otorrhea at visits 3 or 4 yielded changes from CC_e treatment success to treatment failure for 28 sham participants and 24 active treatment participants. Almost all of these 52 changes were due to the presence of otalgia on the day of visit 3. The 52 changes resulted in noticeably lower success rates for both arms.

Secondary Efficacy Endpoints

In this section, the secondary efficacy endpoints are discussed. Detailed results for clinical cure secondary endpoints for the ITT sample are given in Table 5, and detailed results for clinical cure secondary endpoints for the Mic-ITT sample are given in Table 6. In these tables, three versions of each secondary endpoint are examined: CC, the sponsor's version; CC_{ot}, the sponsor's version augmented by taking into account the presence of otorrhea; and CC_b, which also takes into account the presence of otorrhea but fills in missing endpoint values via WCS rather than using the algorithm given in the SAP. As noted above, CC_{ot} is deemed the most clinically relevant version of the clinical cure endpoint by the FDA clinical review staff. CC_b is "worst case" with respect to CC_{ot}: both CC_b and CC_{ot} take into account the presence of otorrhea but they use different policies (WCS vs. SAP algorithm, respectively) for filling in missing endpoint values.

The results of testing the CC_{ot} version of the secondary clinical cure endpoints, in the order in which they are specified in the gatekeeping procedure for controlling FWER, follow. Per Tables 5 and 6, for all endpoints, significant test results were obtained using CC_{ot} if and only if they were also obtained using CC and CC_b.

Clinical Cure at Study Day 15/Visit 4 on the ITT Sample

Per Table 5, for CC_{ot} at day 15/visit 4, the active treatment is superior to sham at $p<.001$ over the ITT sample.

Clinical Cure at Study Day 8/Visit 3 on the Mic-ITT Sample

Per Table 6, for CC_{ot} at day 8/visit 3, the active treatment is superior to sham at $p<.05$ over the Mic-ITT sample.

Clinical Cure at Study Day 15/Visit 4 on the Mic-ITT Sample

Per Table 6, for CC_{ot} at day 15/visit 4, the active treatment is superior to sham at $p<.001$ over the Mic-ITT sample.

Clinical Cure at Study Day 4/Visit 2 on the Mic-ITT Sample

Per Table 6, for CC_{ot} at day 4/visit 2, the active treatment is not shown to be superior to sham at 2-sided alpha = .05 over the Mic-ITT sample.

Note that because of this statistically nonsignificant result, the gatekeeping procedure mandates that analyses of the remaining secondary endpoints (clinical cure at day 4 on the ITT sample and time to cessation of otalgia) are to be considered exploratory analyses.

Clinical Cure at Study Day 4/Visit 2 on the ITT Sample

Per Table 5, for CC_{ot} at day 4/visit 2, the active treatment is superior to sham at $p<.05$ over the ITT sample.

Further details are given in Tables 5 and 6.

Table 5: Results for Secondary Efficacy Endpoints (Clinical Cure at Days 4 and 15) and Related Sensitivity Endpoints for the ITT Sample

Version of Clinical Cure at Study Day	**Proportion of Successes in Active Tx**	**Proportion of Successes in Sham Tx**	**Difference in Proportions**
Day 4/Visit 2 Results			
Sponsor's Version: No Edema/Erythema/ Tenderness w/o Antibiotics (CC)	45.4% (36.6, 54.3)	30.8% (23.0, 39.5)	14.6% (2.3, 26.6) p=.021
Sponsor's Version + No Otorrhea (CC_{ot})[1]	43.1% (34.4, 52.0)	29.2% (21.6, 37.8)	13.8% (1.6, 26.0) P=.028
"Worst Case" + No Otorrhea (CC_b)[1]	43.1% (34.4, 52.0)	30.0% (22.3, 38.7)	13.1% (1.1, 25.3) p=.039
Day 15/Visit 4 Results			
Sponsor's Version: No Edema/Erythema/ Tenderness w/o Antibiotics (CC)	74.6% (66.2, 81.8)	53.1% (44.1, 61.9)	21.5% (9.5, 33.5) p=.00046
Sponsor's Version + No Otorrhea (CC_{ot})[2]	73.8% (65.4, 81.2)	50.8% (41.9, 59.6)	23.1% (11.4, 35.0) p=.0002
"Worst Case" + No Otorrhea (CC_b)[2]	71.5% (63.0, 79.1)	51.5% (42.6, 60.4)	20.0% (8.3, 32.0) p=.0014

Notes. The active treatment arm contained 130 participants and the sham arm also contained 130 participants. Each cell in the table contains a point estimate and a 95% confidence interval (CI). CIs for proportion of successes are computed via the Clopper-Pearson method, and for the difference in proportions unconditional CIs for two independent binomial observations are constructed using the R statistical software (see footnote 7). The cells for difference in proportions also include the 2-sided p-value from a Fisher's exact test, using alpha = .05. CC_{ot} and CC_b are defined in the text above.

8 participants in the active treatment arm and 1 in the sham arm had missing visit 2 data.

9 participants in the active treatment arm and 3 in the sham arm had missing visit 4 data.

[1] 3 active treatment participants and 2 sham participants were treatment failures solely due to the presence of otorrhea.

[2] 1 active treatment participant and 3 sham participants were treatment failures solely due to the presence of otorrhea.

Table 6: Results for Secondary Efficacy Endpoints (Clinical Cure at Days 4, 8, and 15) and Related Sensitivity Endpoints for the Mic-ITT Sample

Version of Clinical Cure at Study Day	**Proportion of Successes in Active Tx**	**Proportion of Successes in Sham Tx**	**Difference in Proportions**
Day 4/Visit 2 Results			
Sponsor's Version: No Edema/Erythema/ Tenderness w/o Antibiotics (CC)	34.6% (22.0, 49.1)	26.8% (15.8, 40.3)	7.8% (-10.9, 26.6) p=.409
Sponsor's Version + No Otorrhea (CC_{ot})[1]	32.7% (20.3, 47.1)	23.2% (13.0, 36.4)	9.5% (-9.4, 28.2) p=.291
"Worst Case" + No Otorrhea (CC_b)[1]	32.7% (20.3, 47.1)	23.2% (13.0, 36.4)	9.5% (-9.4, 28.2) p=.291
Day 8/Visit 3 Results			
Sponsor's Version: No Edema/Erythema/ Tenderness w/o Antibiotics (CC)	63.5% (49.0, 76.4)	35.7% (23.4, 49.6)	27.7% (8.7, 45.3) p=.0067
Sponsor's Version + No Otorrhea (CC_{ot})[2]	59.6% (45.1, 73.0)	33.9% (21.8, 47.8)	25.7% (6.6, 43.3) p=.0117
"Worst Case" + No Otorrhea (CC_b)[2]	59.6% (45.1, 73.0)	33.9% (21.8, 47.8)	25.7% (6.6, 43.3) p=.0117
Day 15/Visit 4 Results			
Sponsor's Version: No Edema/Erythema/ Tenderness w/o Antibiotics (CC)	69.2% (54.9, 81.3)	37.5% (24.9, 51.5)	31.7% (13.0, 49.0) p=.0011
Sponsor's Version + No Otorrhea (CC_{ot})[3]	67.3% (52.9, 79.7)	32.1% (20.3, 46.0)	35.2% (16.4, 52.1) p=.0005
"Worst Case" + No Otorrhea (CC_b)[3]	65.4% (50.9, 78.0)	33.9% (21.8, 47.8)	31.5% (12.5, 48.7) p=.0019

Notes. The active treatment arm contained 52 participants and the sham arm contained 56 participants. Each cell in the table contains a point estimate and a 95% confidence interval (CI). CIs for proportion of successes are computed via the Clopper-Pearson method, and for the difference in proportions unconditional CIs for two independent binomial observations are constructed using the R statistical software (see footnote 7). The cells for difference in proportions also include the 2-sided p-value from a Fisher's exact test, using alpha = .05. CC_{ot} and CC_b are defined in the text above.

2 participants in the active treatment arm and 0 in the sham arm had missing visit 2 data.
No participants had missing visit 3 data.
3 participants in the active treatment arm and 1 in the sham arm had missing visit 4 data.
[1] 1 active treatment participant and 2 sham participants were treatment failures solely due to the presence of otorrhea.
[2] 2 active treatment participants and 1 sham participant were treatment failures solely due to the presence of otorrhea.
[3] 1 active treatment participants and 3 sham participants were treatment failures solely due to the presence of otorrhea.

Time to Cessation of Otalgia

Given the deficiencies in the sponsor's definition and handling of TTCO, we created a new time-to-cessation endpoint, denoted $TTCO_a$. Before defining it, we note that its analysis set consists of participants who (i) recorded a pain score on study day 1, and (ii) either recorded a non-zero pain score on that day or were using analgesics on that day. Effectively, these are the participants who are known to have experienced otalgia at baseline. There were 242 ITT participants in this analysis set, 121 participants in each arm.

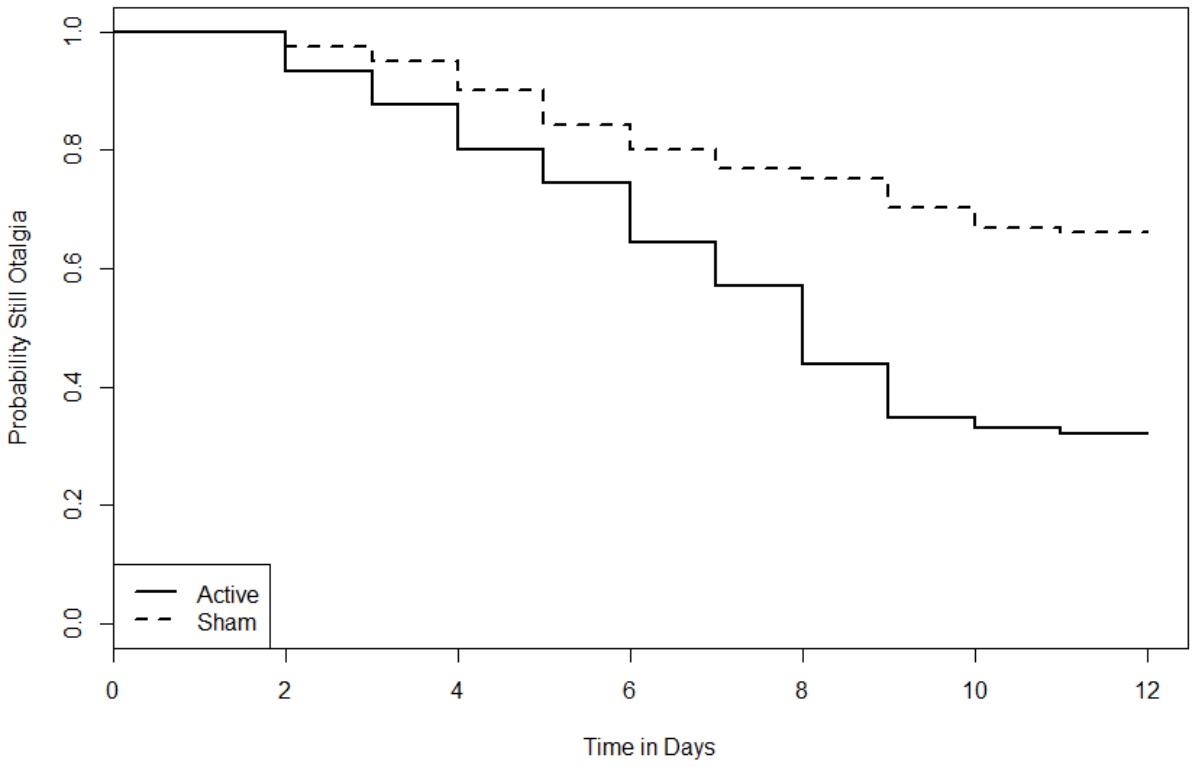
$TTCO_a$ is defined as follows:

1. If the participant used concomitant antibiotics after entry into the study, then $TTCO_a$ is censored at 18 days.
2. Otherwise, if the participant's pain diary ended in a sequence of at least two days with pain scores of 0 and analgesics were not used during the sequence, then $TTCO_a$ is set equal to the first day of the sequence. The requirement of at least two days is intended to provide some proof that the cessation of pain is enduring rather than fleeting. For example, if a participant maintained the diary for 12 days, recorded pain on the first 5 days, recorded no pain on the last 7 days, and last used analgesics on day 8, then $TTCO_a = 9$.
3. Otherwise, $TTCO_a$ is censored at the final study day for which there is a recorded pain score.

One subsequent modification was made to $TTCO_a$. Not all participants maintained the pain diary for the intended 15 or 17 days. For example, 11 participants maintained the diary for 12 days and 20 participants maintained the diary for 13 days. Only 4 participants maintained the diary for fewer than 12 days, however, and we decided to retrospectively administratively censor the $TTCO_a$ data at 12 days in order to avoid having to deal with the possibility of informative censoring. That is, because 238 out of 242 participants in our analysis set maintained the diary for at least 12 days, only analyzing what happens during the first 12 study days effectively avoids having to confront possible estimation biases induced by differential within-arm informative censoring.

In the active treatment arm, 82 participants experienced a cessation of otalgia before 12 days and 41 sham arm participants experienced cessation before 12 days. A log rank test of the difference in the two arm's survival curves yielded chi-square(1) = 27.2, $p < .001$. Figure 1 contains Kaplan-Meier plots of the survival curves.

Figure 1: Kaplan-Meier Curves for Cessation of Otalgia, by Arm (121 Participants in Each)



In summary, through the first 12 days post-treatment, the active treatment arm is significantly superior to the sham arm regarding time to cessation of otalgia. These results are qualitatively similar to the TTCO results reported in the CSR.

3.3 Evaluation of Safety

The safety sample consists of all participants who received study drug. Of the 259 members of the safety sample, 257 also belonged to the ITT sample. 127 of the safety sample members were in the active treatment arm and 132 in the sham arm.

82 adverse events were experienced during the course of the trial, by 54 participants. All of these participants also belonged to the ITT sample. One participant experienced 6 adverse events, one experienced 4 adverse events, five experienced 3 adverse events, ten experienced 2 adverse events, and 37 experienced a single adverse event. 35 of the adverse events were experienced by participants in the active treatment arm, and 47 were experienced by participants in the sham arm. No deaths were reported, and there was one severe adverse event.

Table 7 gives an overview of the adverse events. Almost two thirds of these events were mild, 87% were deemed non-treatment related, and 63% affected the study or other ear. In the active treatment arm, adverse events were slightly more likely to be severe or moderate than in the sham arm (37% vs. 32%), more likely to be probably or possibly treatment related (20% vs. 9%), and less likely to involve the study or other ear (54% vs. 70%).

Table 7: Overview of Adverse Events

Variable	Value	Active Treatment	Sham	Total
Adverse Event	Yes	24(35): 19%	30(47): 23%	54(82): 21%
Severity	Severe	1(1): 1%	0(0): 0%	1(1): <1%
	Moderate	8(12): 6%	12(15): 9%	20(27): 8%
	Mild	17(22): 13%	21(32): 16%	38(54): 15%
Treatment Related?	Probable	0(0): 0%	1(1): 1%	1(1): <1%
	Possible	5(7): 4%	3(3): 2%	8(10): 3%
	Not Related	20(28): 16%	28(43): 21%	48(71): 19%
Affects Ear?	Study Ear	12(13): 9%	14(18): 11%	26(31): 10%
	Non-study Ear	6(6): 5%	10(15): 8%	16(21): 6%
	No	13(16): 10%	13(14): 10%	26(30): 10%

Notes. Cells in columns 3-5 give number of participants (number of adverse events): % of participants in arm or sample. For example, 12 participants in the active treatment arm experienced adverse events related to the study ear, and these participants collectively experienced 13 such events. These 12 participants constitute 9% of the 127 safety sample participants in the active treatment arm. The associated numbers of participants and adverse events differ when at least one participant belonging to a cell experiences more than one adverse event. Also note that some participants are represented in multiple cells defined with respect to a single variable. For example, 24 active treatment participants experienced at least one adverse event, but the sum of the active treatment participants in the severity cells equals 1+8+17 = 26. This is due to the fact that some active treatment participants experienced multiple adverse events of differing levels of severity.

The clinical reviewer provides an in-depth medically-informed analysis of the adverse events.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

This section presents findings for the primary efficacy endpoint regarding important subpopulations. Findings for both the sponsor's version of the endpoint and the preferred version of the endpoint are presented.

The demographic subpopulations examined are (separately) based on gender, race/ethnicity, and age. Geographic region is not examined because all sites are in the United States. In addition, subpopulations based on the presence or absence of a positive culture for *S. aureus* or *P. aeruginosa* at study entry are examined.

4.1 Gender, Race, and Age

Table 8 provides descriptive statistics for the within-arm cure rates and difference in cure rates for day 8 CC and preferred endpoint CC_{ot} for gender, age, and race/ethnicity subgroups. No descriptive statistics were generated for regional subgroups, because all 37 sites were located in the United States.

Table 8: Day 8/Visit 3 Clinical Cure Results by Demographic Subgroup within ITT Sample

Variable	Subgroup	N N_{AT}/N_{Sh}	CC Rates %			CC_{ot} Cure Rates %		
			AT	Sh	Diff	AT	Sh	Diff
Gender	Female	150 (57.7%) 75/75	69.3	48.0	21.3	69.3	45.3	24.0
	Male	110 (42.3%) 55/55	72.7	49.1	23.6	69.1	47.3	21.8
Age	preteen (<13)	42 (16.2%) 17/25	88.2	44.0	44.2	88.2	40.0	48.2
	teen (13-17)	25 (9.6%) 12/13	50.0	30.8	19.2	41.7	30.8	10.9
	adult (18-64)	171 (65.8%) 92/79	72.8	54.4	18.4	72.8	53.2	19.7
	senior (65+)	22 (8.5%) 9/13	44.4	38.5	6.0	33.3	30.8	2.6
Race/ Ethnicity	White & Non-Latino	139 (53.5%) 72/67	68.1	43.3	24.8	66.7	38.8	27.9
	Latino	84 (32.3%) 41/43	70.7	58.1	12.6	68.3	58.1	10.2
	Black & Non-Latino	29 (11.2%) 14/15	92.9	46.7	46.2	92.9	46.7	46.2
	Other	8 (3.1%) 3/5	33.3	40.0	-6.7	33.3	40.0	-6.7

Notes. N = number of members of ITT sample in subgroup. CC = Clinical Cure at day 8 as defined by sponsor. CC_{ot} is defined in the text above and takes into account presence of otorrhea at day 8. AT = active treatment arm. Sh = Sham arm. Diff = difference in cure rates. 130 participants in each arm.

The following observations are based on the raw data given in the table above and not on determinations of statistically significant differences: CC and CC_{ot} yield very similar results regarding treatment effects within subgroups. Regarding gender, females and males show similar results. Within age groups, the treatment effect was by far the largest for preteens and was the smallest for seniors. Within race-ethnicity groups, the treatment effect was largest for blacks and smallest for "Others."

In addition, likelihood ratio tests of arm-by-subgroup interactions were performed to formally examine whether treatment effect magnitude differed within subgroups, separately with regard to CC and with regard to CC_{ot}. These tests were conducted by comparing a logistic regression

model containing main effects for arm and binary subgroup indicator to a model additionally containing an arm-by-subgroup interaction term. The arm-by-gender interaction was not significant, p>.80 for both endpoints. The arm-by-whether-18+ interaction was not significant, p>.30 for both endpoints. The arm-by-whether-white interaction was not significant, p>.50 for both endpoints.

4.2 Other Special/Subgroup Populations

We also examined the treatment effects for CC and CC_{ot} within and external to the Mic-ITT sample. See Table 9. Table 6 above gave the results for the Mic-ITT sample, and they are repeated here.

Table 9: Day 8/Visit 3 Clinical Cure Results by Mic-ITT Sample Membership

Variable	Subgroup	N N_{AT}/N_{Sh}	CC Rates %			CC_{ot} Cure Rates %		
			AT	Sh	Diff	AT	Sh	Diff
In Mic-ITT?	Yes	108 52/56	63.5	35.7	27.7	59.6	33.9	25.7
	No	152 78/74	75.6	58.1	17.5	75.6	55.4	20.2

Notes. N = number of members of ITT sample in subgroup. CC = Clinical Cure at day 8 as defined by sponsor. CC_{ot} is defined in the text above and takes into account presence of otorrhea. AT = active treatment arm. Sh = Sham arm. Diff = difference in cure rates.

Results are similar for both endpoints. Based on the raw data in the table, the Mic-ITT sample had a somewhat larger treatment effect than the non-Mic-ITT sample. However, a likelihood ratio test of the arm-by-subgroup interaction was not significant, p>.50 for both endpoints.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

The reviewer's comments in subsection 3.2 address questionable statistical approaches to handling missing data and to analyzing time-to-cessation-of-otalgia data. Regarding handling missing data, a "worst case scenario" (WCS) approach was utilized in defining alternative clinical cure endpoints. For the primary efficacy endpoint of clinical cure at 8 days/visit 3, using WCS turned out to be equivalent to using the sponsor's missing data imputation algorithm, as given in the SAP. For the secondary clinical cure endpoints, the two missing data approaches diverged somewhat in filling in missing values, but it was always the case that analysis of the sponsor's version of the endpoint showed a statistically significant OTIPRIO superiority if and only if analysis of the WCS-alternative version showed a statistically significant OTIPRIO superiority. Hence, the sponsor's imperfect handling of missing data does not result in any incorrect inferences about the primary and secondary clinical cure endpoints.

Regarding the sponsor's deficiencies in conducting time-to-cessation-of-otalgia analysis, the more accurate analysis presented above yielded qualitatively similar results to those obtained by the sponsor, namely, that OTIPRIO is superior to no treatment. Here, again, the sponsor's analysis yields qualitatively correct inference.

In summary, there were no statistical issues that cloud the evaluation of efficacy.

5.2 Collective Evidence

As noted above, OTIPRIO received FDA approval for the treatment of pediatric patients with bilateral otitis media with effusion undergoing tympanostomy tube placement in December 2015. This approval was based on the review of efficacy and safety data from two Phase 3 randomized multicenter sham-controlled clinical trials that included 532 pediatric patients. When, in IND110244, the sponsor subsequently proposed the additional use of OTIPRIO for the treatment of AOE, the FDA agreed with this additional indication. The FDA also agreed with the sponsor that a single Phase 3 trial would suffice to demonstrate the efficacy and safety of OTIPRIO for the new indication, in part basing this on the demonstration of safety in the Phase 3 trials for bilateral otitis media with effusion. This review has been based entirely on the efficacy and safety data from the single Phase 3 trial for AOE, named 201-201609.

5.3 Conclusions and Recommendations

Analyses of the primary and secondary efficacy endpoints provide very strong evidence of the superiority of OTIPRIO vs. no treatment at days 8 and 15. This superiority is demonstrated across different definitions of clinical cure and different approaches to handling missing data. In addition, examination of safety data did not raise concerns about the safety of the medication. Therefore, approval of OTIPRIO for the treatment of AOE is recommended.

5.4 Labeling Recommendations

The following is the recommended table of clinical response results (with explanatory text) for the label:

One randomized multicenter, sham-controlled clinical trial in 262 pediatric and adult patients with unilateral or bilateral acute otitis externa evaluated the safety and efficacy of OTIPRIO when administered by a healthcare professional as a single dose to the external ear canal. Clinical response was defined as the complete absence of signs and symptoms of acute otitis externa (i.e., tenderness, erythema, edema, and otorrhea), and no concomitant systemic or topical antibiotic (given in the study ear) was taken for any reason at or prior to the study visit. The table contains the proportions of patients with clinical response at Day 8 in both the intent to treat (ITT) population which contains all subjects who were randomized and did not have Group A Streptococci cultured on Day 1 and the microbiological ITT population that contains all ITT subjects who had a positive culture for *S. aureus* or *P. aeruginosa* on Day 1.

Table 10: Proportion of Patients with Clinical Response at Study Day 8 (Acute Otitis Externa)

Study Population	**OTIPRIO**	**Sham**	**% Difference (OTIPRIO - Sham) (95% CI)**
Intention to Treat (ITT) N=260	69% 90/130	46% 60/130	23.1[1] (10.66, 34.62)
Microbiological ITT (Mic-ITT) N=108	60% 31/52	34% 19/56	25.7[2] (6.57, 43.32)

1 $p<0.001$ from a Fisher's exact test.
2 $p=0.012$ from a Fisher's exact test.
ITT population = all subjects who were randomized and did not have Group A Streptococci cultured on Day 1.
Mic-ITT population = all ITT subjects who had a positive culture for *S. aureus* or *P. aeruginosa* on Day 1.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

EDWARD D BEIN
01/23/2018

KAREN M HIGGINS
01/23/2018
I concur.