CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

205434Orig1s000

STATISTICAL REVIEW(S)

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 #:	NDA 205,434
Drug Name:	Flonase (Fluticasone propionate aqueous nasal spray)
Indication(s):	Relief of both the nasal and ocular symptoms associated with seasonal (SAR) and perennial rhinitis (PAR) (b) (4)
Applicant:	Glaxo Smith Kline
Date(s):	9/23/13 (received)
Review Priority:	Standard
Biometrics Division:	DBIV
Statistical Reviewer:	Scott Komo, DrPH
Concurring Reviewers:	Karen Higgins, ScD
Medical Division:	Division of Nonprescription Drug Clinical Evaluation
Clinical Team:	James Stansbury, PhD Steven Osborne, MD Narayan Nair, MD
Project Manager:	Jung Lee, RPh

Keywords: Labeling

Table of Contents

APPEARS THIS WAY ON ORIGINAL

1 EXECUTIVE SUMMARY

An application was submitted for Flonase Allergy Relief (Fluticasone propionate) nasal spray from prescription to over-the-counter (OTC) status for the relief of both the nasal and ocular symptoms associated with seasonal (SAR) and perennial rhinitis (PAR) (b) (4) . Because consumers are unlikely to understand these terms, the Applicant proposed the following language for the Drug Facts Label (DFL): Temporarily relieves the following symptoms due hay fever, other respiratory allergie (b) (4) nasal congestion, runny nose, sneezing, itchy nose, and itchy, watery eyes (b) (4)

The application contained six consumer behavior studies. These included a pilot label comprehension study (RH01305), a targeted label comprehension study (RH01318), a targeted self-selection study (RH01442), two human factor studies (RH01801 and RH01929), and a legacy self-selection / actual use study (R1810198). These studies were reviewed in the NDA documents with the exception of the legacy self-selection/actual use study that was not reviewed because it used an older version of the label and was included in the submission by the Applicant for the safety data it provided.

The pilot label comprehension study (RH01305) enrolled one hundred thirty-three adult subjects from four sites. The primary objectives of the study were to assess consumers' comprehension of the uses, warnings, and direction sections of the DFL. The targeted label comprehension study (RH01318) enrolled six hundred seventeen subjects from nine sites. The primary objectives of the study were to assess consumers' comprehension of selected warnings and directions from the DFL.

The targeted self-selection study (RH01442) was conducted because of concerns of potential drug-drug interactions with other CYP 450 inhibitors, such as ritonavir. Concomitant use of fluticasone propionate and ritonavir can place consumers at increased risk of systemic glucocorticoid side-effects and should be avoided. Four hundred one adult HIV patients prescribed ritonavir were enrolled from nine HIV clinics and clinical sites. The primary objective of the study was to demonstrate that subjects taking ritonavir correctly select not to use fluticasone propionate.

The Applicant conducted two human factor studies. The first study (RH01801) was conducted in forty general population adult subjects enrolled at a single site. The second study (RH01929) was conducted in fifteen low literate adult subjects enrolled at a single site. The primary objectives of both human factors studies were to evaluate the use of fluticasone propionate in an OTC environment, specifically to understand the consumer's ability to clean and prime the nasal spray apparatus correctly, and to demonstrate that the consumer understands how to use the nasal spray, including the correct route of administration of the product (intranasal versus intraocular).

In the label comprehension studies (RH01305 and RH01318), the uses, directions, and warnings were generally well understood by both the general and low literate populations.

In the targeted self-selection study of HIV subjects taking ritonavir (RH01442), less than half of the adult HIV patients correctly selected not to use the product. This was much lower than the 95% performance threshold. The low correct self-selection rate occurred regardless of literacy level, history of nasal allergies, and previous use of fluticasone propionate. The low correct self-selection rate is concerning because concomitant use of fluticasone propionate and ritonavir could place consumers at increased risk for systemic glucocorticoid side-effects. The Applicant proposed modifications to the DFL that would likely improve the correct self-selection rate but it is impossible to estimate the magnitude of the likely improvement as an additional self-selection study was not conducted using the modified label.

In the human factors studies (RH01801 and RH01929), all of the subjects in both studies correctly understood the correct route of administration (intranasal vs. intraocular). For the Prime and Wipe/Rinse tasks, correct task performance fell well short of the 80% success criterion for both the general and low literate population at both the "Initial Use" and "Two Week Later" sessions. Subjects who failed a task at either session were directed to review the package insert (Quick Start Guide) and repeat the task. During this repeat sessions, task performance for both general and low literate population subjects improved greatly. The Applicant viewed the Repeat session results as an assessment of the Quick Start Guide's ability to aid consumer understanding of correct product use. Because many subjects did not initially review the Quick Start Guide and review of the Quick Start Guide increased correct performance for many of the tasks, the Applicant proposed changes to the packaging to increase review of the Quick Start Guide.

In both the general and low literate populations, subjects exhibited poor performance for the task of cleaning after storing the product for at least one week at both testing sessions. During the "Repeat" session for subjects who made errors during either the "Initial Use" or "Two Week Later" session, performance improved but still fell far below the 80% success criterion with subjects having the most difficulty with the sub-step that asked subjects what they would do if the nozzle became clogged. The results for "Repeat" session suggest that the changes to the packaging to increase consumer review of the Quick Start Guide will likely increase the rate of proper cleaning but does not provide confidence that it would meet the 80% success criterion.

Overall, subjects in the general and low literate populations generally understood the warnings and directions in the label. There were issues with subject's understanding that they should not take fluticasone propionate and ritonavir together. The Applicant proposed labeling changes to aid consumer understanding that these drugs should not be co-administered. The labeling changes will likely increase correct self-selection but it is difficult to estimate the magnitude of likely correct self-selection improvement as the modified label was not studied in another self-selection study. The determination of the adequacy of the proposed label modification to minimize concomitant use of fluticasone propionate and ritonavir is left to the clinical reviewer as this decision requires taking into account the safety risk of taking both drugs together.

Subjects had some difficulty in correctly cleaning the product after storing it for at least one week. Most errors occurred in the sub-step where subjects stated how they would unclog the device if a clog occurred. The determination of whether errors in cleaning resulting in clogged devices would be considered a safety risk is left to the clinical reviewer.

2 INTRODUCTION

2.1 Overview

An application was submitted for Flonase Allergy Relief (fluticasone propionate) nasal spray to switch from prescription to over-the-counter (OTC) status for the relief of both the nasal and ocular symptoms associated with seasonal (SAR) and perennial rhinitis (PAR) (b) (4) in adults. Because consumers are unlikely to understand these terms, the Applicant proposed the following language for the Drug Facts Label (DFL): Temporarily relieves the following symptoms due hay fever, other respiratory allergies (b) (4) (b) (4) nasal congestion, runny nose, sneezing, itchy nose, and itchy, watery eyes in years and older.

Fluticasone propionate is a second-generation corticosteroid. It is a synthetic trifluorinated corticosteroid (S-(fluoromethyl) 6α, 9-difluoro-11β-17-dihydroxy-16α-methyl- 3-oxoandrosta-1,4-diene-17β-carbothioate, 17-propionate).

Each actuation of the nasal spray provides 50 mcg of fluticasone propionate. The recommended dose in adults is 2 sprays into each nostril (200 mcg total) once daily (QD) in the first week and 1 or 2 sprays into each nostril QD in week 2 and onward.

Regulatory History
Fluticasone propionate was originally approved in 1994 for the treatment of nasal symptoms of seasonal allergic rhinitis (SAR) and perennial allergic rhinitis (PAR) in adults and children 12 years of age and older. Fluticasone propionate was subsequently approved in 1997 for pediatric patients 4 years of age and older. In 1998, it was approved in patients with perennial non-allergic rhinitis. Finally, it was approved in 2002 for as-needed (PRN) use.

The Agency, at a 2001 meeting, had several recommendations on the label comprehension study that comprehension should be tested on how long to use Flonase, any medical conditions that might contraindicate use of Flonase, and potential drug-drug interactions with CYP 450 inhibitors. In addition, the Agency also recommended that an actual use study be conducted.

Between 2001 and 2003, the Applicant conducted a series of label comprehension studies, self-selection studies and an actual use trial (AUT). In 2004, for business reasons not related to the program itself, the Applicant decided to put the Rx-to-OTC switch program on hold.

In 2011, the Applicant decided to reactivate the Rx-to-OTC switch program. The Drug Facts Label (DFL), developed from 2001–2003, used in the previous Actual Use Trial was revised to account for changes in clinical practices and other OTC allergy treatments. The Applicant included the AUT in the submission, even though the label changed over time, because it contained safety information in a naturalistic setting without the intervention of health professional for up to six months.

At the February 2011 meeting, the Applicant and FDA discussed label comprehension (LC), self-selection (SS) and actual use (AU) studies. The Applicant proposed use of the product in

(b) (4) years of age and older. (b) (4) The Applicant proposed two new Actual Use studies, one lasting 2 weeks (all comers with nasal allergies) and one lasting 4 months (2 groups, one who never used Flonase and one that had used Flonase previously). Of note, FDA recommended an actual use study of 6 months duration, to allow a full assessment of usage patterns.

In an April 2012 advice letter, the Agency stated the following:

> *As noted in the meeting minutes for the February 22, 2011 Type B meeting, we find the proposed labeling* (b) (4)

At an October 2012 meeting, the Agency stated their concerns with the proposal to add ocular symptoms to the label. The Agency stated that the Applicant would need to demonstrate in label comprehension that consumers understand that this product is to be used intranasally and not intraocularly. In addition, data would also be needed to show that consumers do not use Flonase in their eyes. The Agency said that it was unclear whether an actual use component was needed. The Agency advised the Applicant to submit the human factors study protocol for review and comment.

The Applicant did not initiate, and Agency later dropped the requirement for, a second actual use study. The Agency decided consumers understood the concept of a nasal spray drug and a competing corticosteroid nasal spray drug (triamcinolone acetonide) was not required to perform an actual use study.

The Applicant conducted a human factors study (RH01801) in a general population. The Agency recommended that the Applicant conduct a human factors study in at least 15 low literate adults (May 2013 advice letter). Literacy was not prospectively measured in Study RH01801 so the Applicant conducted a human factors study in a low literacy population (Study RH01929).

2.2 Data Sources

The consumer studies submitted in the application are presented in the following table.

Table 1: Consumer behavior studies reviewed in this application

Applicant defined study number	*Study Type*	*# of Subjects*
RH01305	Pilot label comprehension	130
RH01318	Targeted label comprehension	607
RH01442	Targeted self-selection	399
RH01801	Human factors	40
RH01929	Human factors in a low literacy population	15

Note: The application also contained a self-selection/actual use study (R1810198) that was not reviewed because it used an older version of the label. The study was included in the Application for the safety data it provided.

Clinical study reports are available at:

\\CDSESUB1\evsprod\NDA205434\0000\m1

Datasets are available at:

\\CDSESUB1\evsprod\NDA205434\0000\m5

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The submission contained analysis datasets along with a define.pdf and an annotated CRF for the five consumer behavior studies reviewed in this application. I was able to reproduce the primary and secondary analyses for all of the studies. It should be noted that for Studies RH01801 and RH01929, the study report and protocol stated that confidence intervals would be calculated using exact methods. However, the confidence intervals presented in the reports for Studies RH01801 and RH01929 were based on normal approximations. The confidence intervals presented in this review for Studies RH1801 and RH01929 are computed using the Clopper-Pearson method to construct exact confidence intervals.

3.2 Evaluation of Efficacy

The labeling was developed in an iterative fashion based on the results on the individual studies. The version of the label used in studies is included in the Appendix. In addition, the final labeling is also included in the Appendix.
The results will be presented for the consumer behavior studies presented in Table 1. The pilot label comprehension study (RH01305) will be discussed in Section 3.2.1, the targeted label comprehension study (RH01318) will be discussed in Section 3.2.2, the targeted self-selection study (RH01442) will be discussed in Section 3.2.3, the human factors study (RH01801) will be discussed in Section 3.2.4, and the human factors study in a low literate population (RH01929) will be discussed in Section 3.2.5.

3.2.1 Pilot Label Comprehension Study (RH01305)

3.2.1.1 Study Design and Endpoints (RH01305)

The Applicant conducted a pilot label comprehension study in adults who self-reported having nasal allergies within the last year. Subjects were given the Rapid Estimate of Adult Literacy in Medicine (REALM) to measure health literacy. Subjects were given the draft labeling and then were asked the label comprehension questions. Each question was followed up with the question, "Why do you say that".

One hundred thirty-three subjects were enrolled from four sites, of which one was devoted solely to recruiting low literate subjects (See Table 2). There were one hundred thirty eligible subjects who were recruited into two cohorts. Cohort 1, which represented the general population, included one hundred six subjects and Cohort 2, which represented the low literate population defined by a REALM score ≤60, included twenty-four subjects.

Table 2: Study Sites for Study RH01305

Site	Location
1	**INDIANAPOLIS**
2*	**BALTIMORE**
3	**LOS ANGELES**
4	**DALLAS**

* Site 2 was devoted to recruiting only low literate subjects

The primary objective of the study was to measure comprehension of the following label elements:

1. Uses
    a. Relieves symptoms of indoor and outdoor allergies
        i. Sneezing
        ii. Itchy nose
        iii. Runny nose
        iv. Nasal congestion
2. Warnings
    a. Do not use to treat asthma
    b. Ask a doctor for pharmacist before use if you are using ritonavir (medicine for HIV infection)
    c. Stop use and ask a doctor if
        i. Your symptoms do not get better within 7 days of starting use. You may have something more than allergies, such as an infection.
        ii. You get new symptoms such as severe facial pain or thick nasal discharge. You may have something more than allergies, such as an infection.
3. Directions
    a. Use this product only once a day
    b. (b) (4) years of age and older, Week 1, use 2 sprays in each nostril once daily
    c. (b) (4) years of age and older, Week 2 onwards, use 1 or 2 sprays in each nostril once daily, as needed to treat your symptoms

d. (b) (4) years of age and older, After (b) (4) months of daily use, ask your doctor if you can keep using

Label comprehension endpoints

To assess label comprehension, subjects were read scenarios and asked to provide a response to assess their comprehension. For each primary objective, the scenario is listed in italics.

- Relieves symptoms of indoor and outdoor allergies: sneezing, itchy nose, runny nose, and nasal congestion.

  According to the label, what is this product used for?

- Do not use to treat asthma

  Bob has asthma symptoms and lost his asthma medication. He would like to use Flonase to treat his asthma symptoms. According to the product label, is it okay or not okay for Bob to use the product to treat his asthma?

- Ask a doctor for pharmacist before use if you are using ritonavir (medicine for HIV infection)

  Sally has HIV and currently takes a prescription medication which contains ritonavir. She would like to start taking Flonase to treat her allergies. According to the product label, what, if anything, should Sally do?

- Stop use and ask a doctor if your symptoms do not get better within 7 days of starting use. You may have something more than allergies, such as an infection.

  Vera started using this product 7 days ago for her allergy symptoms. Her symptoms have gotten worse each day. According to the product label, what, if anything, should Vera do?

- Stop use and ask a doctor if you get new symptoms such as severe facial pain or thick nasal discharge. You may have something more than allergies, such as an infection.

  Maria has been using this product for her allergies and it's been working. But, over the last 7 days, she feels worse. She gets severe facial pain when she bends down. She has never experienced this kind of face pain before. She also has a new thick yellow nasal discharge.

- Use this product only once a day

  This morning Lionel used 2 sprays of Flonase in each nostril to treat his allergy symptoms. At bedtime, he notices his symptoms have returned and would like to take another dose. According to the product label, is it okay or not okay for Lionel to take another dose of Flonase?

- (b) (4) years of age and older, Week 1, use 2 sprays in each nostril once daily

  According to the product label, how many sprays can you use in each nostril daily during the first week of treatment?

- (b) (4) years of age and older, Week 2 onwards, use 1 or 2 sprays in each nostril once daily, as needed to treat your symptoms

According to the product label, how many sprays can you use in each nostril daily starting at week 2 of treatment?

- (b) (4) years of age and older, After (b) (4) months of daily use, ask your doctor if you can keep using

  Lewis has been using this product daily for (b) (4) months to treat his nasal allergies. He is considering using Flonase for another month. According to the product label, what, if anything, should Lewis do?

Label comprehension responses were coded in the following way:

1. Initial responses to the scenario question were coded as 'Correct', 'Acceptable', or 'Incorrect', based on pre-defined coding categories.
2. Final comprehension was determined using the initial response along with the follow-up open-ended responses. Responses were categorized as either "Correct" or "Incorrect".

3.2.1.2 Statistical Methodologies (RH01305)

Primary label comprehension endpoints
The primary endpoints were the correct response rate for the nine primary label comprehension communication objectives. The comprehension rate was computed for each communication objective as the number of subjects who provided a correct response divided by the number of subjects who answered the question. In addition, the two-sided 95% exact confidence interval was computed for the correct response rate of each communication objective. Comprehension rates were computed separately for the general population (Cohort 1) and the low literate population (Cohort 2). Because this was a pilot study, the comprehension rates were not compared to a performance threshold.

3.2.1.3 Patient Disposition, Demographic and Baseline Characteristics (RH01305)

Disposition of subjects
One hundred thirty-three subjects were enrolled in the study. Three subjects who were screened from the site that was recruiting only low literate subjects had a REALM score of greater than 60 so they did not complete the interview and were not included in the analyses. To make up for the loss of these subjects, three low-literate subjects were recruited from a different site that was recruiting for the general population.

Subject demographics
As can be seen in Table 3, there were approximately equal numbers of males and females in both the general population and low literates subjects. There were more subjects at least 65 years of age in the general population (20.8%) compared to the low literate subjects (12.9%). In the general population, the most common race was Caucasian/White (68.9%) in contrast to the low literate subjects where the most common race was African American/Black (67.7%). For the general population, very few (<1%) of the subjects did not complete high school. In contrast for the low literate cohort, approximately 2/3 of the subjects did not complete high school. In the general population, approximately ¼ of the subjects used fluticasone propionate previously while approximately 10% of the low literate subjects used fluticasone propionate previously.

Table 3: Demographics for Study RH01305

	Total (N=130)		General Population (N=106)		Low Literate† (N=31)	
	n	%	n	%	n	%
Gender						
Male	62	47.7	50	47.2	16	51.6
Female	68	52.3	56	52.8	15	48.4
Age						
18 to 24	11	8.5	8	7.5	4	12.9
25 to 34	18	13.8	15	14.2	4	12.9
35 to 44	25	19.2	17	16.0	8	25.8
45 to 54	24	18.5	20	18.9	7	22.6
55 to 64	28	21.5	24	22.6	4	12.9
65 or older	24	18.5	22	20.8	4	12.9
Race						
Caucasian/White	80	61.5	73	68.9	10	32.3
African American/Black	37	28.5	20	18.3	21	67.7
Asian or Pacific Islander	4	3.1	4	3.7	0	0.0
Other	9	6.9	9	8.3	0	0.0
Hispanic Origin						
Yes	11	8.5	9	8.5	2	6.5
No	119	91.5	97	91.5	29	93.5
Education						
Less than High School	22	16.9	1	0.9	21	67.7
Completed High School	16	12.3	14	13.2	4	12.9
Some College/Technical	46	35.4	45	42.5	5	16.1
Graduated College / Technical School or More	46	35.4	46	43.4	1	3.2
Income						
$0 to $14,999	27	20.8	7	6.6	22	71.0
$15,000 to $24,999	16	12.3	16	15.1	1	3.2
$25,000 to $34,999	16	12.3	14	13.2	5	16.1
$35,000 to $44,999	17	13.1	15	14.2	3	9.7
$45,000 to $64,999	22	16.9	22	20.8	0	0.0
$65,000 to $74,999	8	6.2	8	7.5	0	0.0
$75,000 or more	17	13.1	17	16.0	0	0.0
Refused	7	5.4	7	6.6	0	0.0
REALM Category						
Low-Literacy	31	23.8	7	6.6	31	100.0
Normal Literacy	99	76.2	99	93.4	0	0.0

Previous Fluticasone Propionate Use						
Yes	27	20.8	26	24.5	3	9.7
No	103	79.2	80	75.5	28	90.3

† Includes low-literate subjects recruited into cohort 1 (general population, n=7) and subjects recruited into cohort 2 (low-literate, n=24)
Source: Study report, Table 4

3.2.1.4 Results and Conclusions (RH01305)

The label comprehension rates for the primary communication objectives are presented in Table 4. Comprehension was relatively high for the majority of the primary communications. However, comprehension of the "Stop use and ask a doctor if your symptoms do not get better within 7 days of starting use" warning was not that well understood (General population: 79.2%, 95% exact lower confidence bound (LCB) = 70.3%; low literate population: 54.8%, 95% exact LCB = 36.0%). Of the 22 subjects with an incorrect response, 19 mentioned "ask a doctor" but failed to mention "stop using the product" in their initial response. Upon follow up, 6 of these 19 subjects mentioned that the medicine was not working or that a condition other than allergies, such as an infection, might be present.

For the "Stop use and ask a doctor if you get new symptoms such as such as severe facial pain or thick nasal discharge" and "Use this product only once a day" communication objectives, the low literate subjects had lower comprehension rates than the general population.

Table 4: Comprehension rates for the Primary Communication Objectives (RH01305)

Primary Communication Objectives	General Population (N=106)			Low Literate† (N=31)		
	n	%	LB*	n	%	LB*
(Q14) (b)(4) years of age and older, Week 1, use 2 sprays in each nostril once a day	106	100.0	96.6	31	100.0	88.8
(Q18) (b)(4) years of age and older, Week 2 onwards, use 1 or 2 sprays in each nostril once a day, as needed to treat your symptoms	106	100.0	96.6	31	100.0	88.8
(Q9) Ask a doctor or pharmacist before use if you are taking ritonavir	105	99.1	94.9	30	96.8	83.3
(Q3) Product indication (Uses)	105	99.1	94.9	30	96.8	83.3
(Q13) (b)(4) years of age and older, after (b)(4) months of daily use, ask your doctor if you can keep using	102	96.2	90.6	30	96.8	83.3
(Q4) Do not use to treat asthma	102	96.2	90.6	29	93.5	78.6
(Q10) Stop use and ask a doctor if you get new symptoms such as such as severe facial pain or thick nasal discharge	101	95.3	89.3	25	80.6	62.5
(Q11) Use this product only once a day	98	92.5	85.7	25	80.6	62.5
(Q7) Stop use and ask a doctor if your symptoms do not get better within 7 days of starting use	84	79.2	70.3	17	54.8	36.0

† Includes low-literate subjects recruited into cohort 1 (general population, n=7) and subjects recruited into cohort 2 (low-literate, n=24).
n: #correct responses
* Lower two-sided 95% exact confidence bound
Source: Study report, Tables 5 and 7

3.2.2 Targeted Label Comprehension Study (RH01318)

Based on the results of the pilot label comprehension study (RH01305), the Applicant made the following changes to the DFL:

- The Directions were modified to reflect (b) (4) years and older and a duration of use of 6 months. It should be noted that the age limit was subsequently (b) (4) years and older.
- The warning "Stop use and ask a doctor if you get changes to your vision (b) (4) " was added to be consistent with the communication of other rare side-effects (b) (4)) included in the label.
- The warning, "Ask a doctor or pharmacist before use if you are taking (b) (4) taking ritonavir (medicine for HIV infection) because of the concerns with co-administration of Flonase and ritonavir.
- The Uses section was modified based on Agency recommendation to (b) (4) modify the language to be consistent with other products used for the treatment of allergic rhinitis, i.e. for the temporary relief of symptoms of hay fever or other respiratory allergies. The Uses section was also revised to add "itchy, watery eyes" to the list of allergic rhinitis symptoms (b) (4)

3.2.2.1 Study Design and Endpoints (RH01318)

The Applicant conducted a targeted comprehension study in adults. Subjects were given the REALM to measure health literacy. Subjects were given the draft labeling and then were asked the label comprehension questions. Each question was followed up with the question "Why do you say that".

Six hundred seventeen subjects were enrolled from nine sites, of which two were devoted solely to recruiting low literate subjects (See Table 5). The six hundred seven subjects were divided into two cohorts. Cohort 1, which represented the general population, included three hundred ten subjects. Cohort 2, which represented the low literate population defined by a REALM score ≤60, included ninety-seven.

Table 5: Study Sites for Study RH01318

Site Number	Location	Site Number	Location
1	Raleigh , NC	**6**	Indianapolis, IN
2	Springfield, Mo	**7**	Tampa, FL
3	Chicago, IL	**8***	Los Angeles, CA
4	Dallas, TX	**9***	Baltimore, MD
5	Seattle, WA		

* Sites 8 and 9 were devoted to recruiting only low literate subjects
Source: Applicant's response to the Agency's 6/9/14 information request

The primary objective of this study was to measure consumer comprehension of the following:

1. Warnings
    a. Stop use and ask a doctor if
        i. Your symptoms do not get better within 7 days of starting use. You may have something more than allergies, such as an infection.
        ii. You get new symptoms such as severe facial pain or thick nasal discharge. You may have something more than allergies, such as an infection.
2. Directions
    a. (b) (4) years of age and older, Week 1, use 2 sprays in each nostril once daily
    b. (b) (4) years of age and older, Week 2 onwards, use 1 or 2 sprays in each nostril once daily, as needed to treat your symptoms.

Label comprehension endpoints
To assess label comprehension, subjects were read scenarios and asked to provide a response to assess their comprehension. For each primary objective, the scenario is listed in italics:

- Stop use and ask a doctor if your symptoms do not get better within 7 days of starting use. You may have something more than allergies, such as an infection.

    Vera started using this product 7 days ago for her allergy symptoms. Her symptoms have not gotten better. According to the product label, what, if anything, should Vera do? Why do you say that?

- Stop use and ask a doctor if you get new symptoms such as severe facial pain or thick nasal discharge. You may have something more than allergies, such as an infection.

    Maria has been using this product for her allergies and it's been working. But lately, she gets severe facial pain when she bends down. She has never experienced this kind of face pain before. She also has a new thick yellow discharge. According to the product label, what, if anything, should Maria do? Why do you say that?

- (b) (4) years of age and older, Week 1, use 2 sprays in each nostril once daily

    According to the product label, how many sprays can you use in each nostril daily during the first week of treatment? Why do you say that?

- (b) (4) years of age and older, Week 2 onwards, use 1 or 2 sprays in each nostril once daily, as needed to treat your symptoms.

    According to the product label, how many sprays can you use in each nostril daily starting at week 2 of treatment? Why do you say that?

3.2.2.2 Statistical Methodologies (RH01318)

Primary label comprehension endpoints
The primary endpoints were the proportion of subjects who gave an overall correct response for the four primary communication objectives in the general population. In addition, a 2-sided 95% exact confidence interval was constructed. If the lower bound of the two-sided 95% confidence

limit was greater than the 90% performance threshold, then it was determined that the primary endpoint met the performance threshold.

The Applicant provided the following clinical justification for the 90% performance thresholds used in this study:

> *The target level of comprehension reflects not only the clinical significance of the primary communication objectives but also the impact of a lower target threshold on the overall execution of the study. The clinical significance of the primary communication objectives described in the paragraphs above indicates that a lower target threshold (i.e., 80-85%) would be appropriate. However, the larger sample size that would be required for a lower target was not warranted, particularly given the balance of the clinical significance and the fact that levels of comprehension demonstrated in pilot testing suggest that the 90% target was achievable.*

The rationale for part of the Applicant's justification that a larger sample size would be required for a lower target threshold is not clear. An increase in sample size would only occur if the assumed comprehension rate used in the sample size calculation was also lower than the one assumed in the present sample size calculation of 94%. Nonetheless, using a higher performance threshold than what is clinically required would only increase the level of comprehension required to demonstrate adequate comprehension of the communication objectives.

Sample size
The sample size for the general population was targeted to be 500 subjects. This sample size provided 90% power to conclude that the true comprehension rate was at least 90% when the observed comprehension rate was 94% or higher using a two-sided exact test at the 5% level. In addition, it was expected that approximately 10% of the general population would have low literacy as defined by a REALM score (< 60). An additional one hundred low literate subjects were recruited to meet the target of one hundred fifty low literate subjects.

3.2.2.3 Patient Disposition, Demographic and Baseline Characteristics (RH01318)

Disposition of Subjects
Of the six hundred seventeen subjects enrolled in the study, five subjects were excluded from the analyses because they did not meet inclusion/exclusion criteria but were interviewed and completed the study. One subject was excluded because it was determined that the interview was not conducted properly. Finally, four subjects who were screened from the site that was recruiting only low literate subjects had a REALM score of greater than 60 so they did not complete the interview and were not included in the analyses. 607 evaluable subjects were included in the analysis. Thus, there were six hundred seven evaluable subjects.

Subjects Demographics
In Table 6, it can be seen that the study included slightly more female than male subjects for both the general population and low literate subjects. In the general population, 16.1% of the subjects were 65 years of age or older. Similarly, 13.1% of the low literate subjects were 65 years of age or older. The most common race in the general population was Caucasian/White (79.6%) and in the low literate subjects, the most common race was African American/Black (57.5%). In the

general population, 56.3% of the subjects experienced nasal allergies within the last year. For low literate subjects, a slightly lower percentage (41.8%) of the subjects experience nasal allergies within the last year. Of the general population subjects who suffered from indoor or outdoor allergies within the last year, 18.8% had used fluticasone propionate within that time. For the low literate subjects who suffered indoor or outdoor allergies during the last year, 7.8% had used fluticasone propionate during that time.

Table 6: Demographics for RH01318

	Total (N=607)		General Population (N=310)		Low Literate† (N=153)	
	n	%	n	%	n	%
Gender						
Male	269	44.3	221	43.3	73	47.7
Female	338	55.7	289	56.7	80	52.3
Age						
18 to 24	64	10.5	53	10.4	18	11.8
25 to 34	84	13.8	74	14.5	23	15.0
35 to 44	107	17.6	97	19.0	18	11.8
45 to 54	138	22.7	107	21.0	41	26.8
55 to 64	120	19.8	97	19.0	33	21.6
65 or older	94	15.5	82	16.1	20	13.1
Race						
Caucasian/White	427	70.3	406	79.6	48	31.4
African American/Black	138	22.7	75	14.7	88	57.5
Native American	2	0.3	1	0.2	1	0.7
Asian or Pacific Islander	12	2.0	9	1.8	3	2.0
Other	28	4.6	19	3.7	13	8.5
Hispanic Origin						
Yes	34	5.6	24	4.7	14	9.2
No	572	94.2	485	95.1	139	90.8
Missing	1	0.2	1	0.2	0	0
Education						
Less than High School	63	10.4	16	3.1	53	34.6
Completed High School	149	24.5	106	20.8	64	41.8
Some College/Technical School	186	30.6	179	35.1	24	15.7
Graduated College/Technical School or More	209	34.4	209	41.0	12	7.8
Income						
$0 to $14,999	132	21.7	63	12.4	82	53.6
$15,000 to $24,999	79	13	67	13.1	22	14.4
$25,000 to $34,999	66	10.9	60	11.8	14	9.2
$35,000 to $44,999	69	11.4	64	12.5	11	7.2

$45,000 to $64,999	77	12.7	72	14.1	12	7.8
$65,000 to $74,999	47	7.7	47	9.2	2	1.3
$75,000 or more	120	19.8	120	23.5	8	5.2
Refused	17	2.8	17	3.3	2	1.3
REALM Category						
Low-Literate	153	25.2	56	11.0	153	100.0
Literate	454	74.8	454	89.0	0	0.0
History (1yr) of Nasal Allergies						
Yes	327	53.9	287	56.3	64	41.8
No	274	45.1	219	42.9	87	56.9
Missing	6	1	4	0.8	2	1.3
Previous (1yr) FP Use*						
Yes	58	17.7	54	18.8	5	7.8
No	269	82.3	233	81.2	59	92.2

† Includes low-literate subjects recruited into cohort 1 (general population, n=56) and subjects recruited into cohort 2 (low-literate, n=97).
*Fluticasone propionate (FP) use in the past 12 months was assessed among all subjects reporting a recent history of indoor or outdoor nasal allergies.
Source: Study report, Table 4

3.2.2.4 Results and Conclusions (RH01318)

The label comprehension rates for the primary communication objectives are presented in Table 7. Although the comprehension of the "Stop use and ask a doctor if your symptoms do not get better within 7 days of starting use" communication objective increased from the pilot label comprehension study, the comprehension rate fell slightly below the 90% performance threshold (General Population: 91.0 (464/510), 95% exact LCB = 88.2%). In the general population, the "Stop use and ask a doctor if you get new symptoms such as severe facial pain or thick nasal discharge" communication objective met the 90% threshold (comprehension rate = 95.1% [485/510], 95% exact LCB = 92.8%). It should be noted for this objective that comprehension in the low literacy population [85.0% (130/153), 95% exact LCB = 78.3%] was lower than the general population. For the two communication objectives related to dosing (Questions 13 and 7), comprehension in the general population met the 90% performance threshold and the low literacy group had similar rates to the general population.

Table 7: Comprehension Rates for Primary Communication Objectives (RH01318)

(Question) Primary Communication Objective	General Population (N=510)			Low Literate † (N=153)		
	n	%	LB*	n	%	LB*
(Q8) Stop use and ask a doctor if your symptoms do not get better within 7 days of starting use	464	91.0	88.2	143	93.5	88.3
(Q6) Stop use and ask a doctor if you get new symptoms such as such as severe facial pain or thick nasal discharge	485	95.1	92.8	130	85.0	78.3

(Q13) (b) (4) years of age and older, Week 1, use 2 sprays in each nostril once a day	509	99.8	98.9	151	98.7	95.4
(Q7) (b) (4) years of age and older, Week 2 onwards, use 1 or 2 sprays in each nostril once a day, as needed to treat your symptoms	509	99.8	98.9	149	97.4	93.4

† Includes low-literate subjects recruited into cohort 1 (general population, n=56) and subjects recruited into cohort 2 (low literate, n=97).
* Lower two-sided 95% exact confidence bound
Source: Study report, Tables 5 and 7

3.2.3 Targeted Self-selection Study (RH01442)

This study was conducted because of concerns of potential drug-drug interactions with other CYP 450 inhibitors, such as ritonavir. Concomitant use of fluticasone propionate and ritonavir can place consumers at increased risk of systemic glucocorticoid side-effects and should be avoided. Because consumers who are taking HIV medication, specifically ritonavir, have the highest potential risk, the self-selection study focused on the label element, (b) (4) if you are taking ritonavir (a medicine for HIV infection)".

3.2.3.1 Study Design and Endpoints (RH01442)

The Applicant conducted a targeted self-selection in adult HIV patients who are prescribed ritonavir. Using medical records, HIV clinics and clinical sites identified subjects who were prescribed ritonavir. These subjects were invited to participate in this study; Four hundred one adult HIV patients were enrolled from six HIV clinics and clinical sites (See Table 8).

Table 8: Study Sites for Study RH01442

Site	Location
1	PHOENIX, AZ
2	SAN FRANCISCO, CA
3	ALLENTOWN, PA
4	BEVERLY HILLS, CA
5	MIAMI BEACH, FL
6	AUSTIN, TX

The primary objective of the study was to demonstrate that subjects taking ritonavir correctly select not to use fluticasone propionate.

Subjects were given the REALM to measure health literacy. Subjects were given the draft labeling and then were asked the self-selection question, "Is it okay to use this product?". They were then asked the follow-up questions, "Why do you say that" and "Anything else".

3.2.3.2 Statistical Methodologies (RH01442)

Primary self-selection endpoint
Self-selection responses were coded in the following way:

1. Responses to the initial self-selection question will be classified as correct ("No") or incorrect ("Yes," "Don't know").

2. Responses to the follow-up questions were coded *a priori* into general coding categories for verbatim responses were:

Follow-up Question Response	Follow-up Response Code
Reference to "do not use": ritonavir/HIV medicine	Correct
Reference to medications they are taking	Acceptable
Reference to having HIV or AIDS	Acceptable
Would ask doctor first/doesn't use medication without talking to doctor	Acceptable
Reference to "ask doctor or pharmacist": glaucoma, ketoconazole use	Acceptable
Reference to "do not use": asthma, unhealed injury/surgery, allergic reaction to product or ingredients	Acceptable
Doesn't have allergies/doesn't treat allergies/uses other allergy products (without reference to any of the correct or acceptable responses indicated above)	Incorrect
Don't know	Incorrect

3. Mitigation was performed using both the initial and follow-up response codes to arrive at a final self-selection response using the following algorithm:

Initial/Follow-up Response Codes	Final Self-selection Code
Correct initially, Correct at follow-up	Correct Overall Response
Correct initially, Acceptable at follow-up	Correct Overall Response
Correct initially, Incorrect at follow-up	Incorrect Overall Response
Incorrect initially, Correct at follow-up	Correct Overall Response
Incorrect initially, Acceptable at follow-up	Correct Overall Response
Incorrect initially, Incorrect at follow-up	Incorrect Overall Response

The correct decision rate will be computed as the number of eligible subjects with a final code of "correct" divided by the number of eligible subjects. A two-sided 95% exact confidence interval for the correct decision rate was also computed.

The correct decision rate was computed separately for subgroups by literacy level, history of nasal allergies, and previous fluticasone propionate nasal spray use.

The study objective was determined to be met if the two-sided 95% exact lower confidence bound for the correct decision rate was at least 95% in the entire sample.

Sample size

The sample size of 390 subjects provided 90% power to conclude that the correct decision rate was at least 95% when the observed comprehension rate was 98% or higher using a two-sided exact test at the 5% level. It was also expected that approximately 10% of the sample would be considered to have low literacy (REALM score <60).

The Applicant stated that 95% performance threshold was selected because co-administration of fluticasone propionate and ritonavir is associated with potentially clinically significant elevated fluticasone propionate levels.

3.2.3.3 Patient Disposition, Demographic and Baseline Characteristics (RH01442)

Disposition of subjects

The study enrolled four hundred one subjects. Two of these subjects had their prescription for ritonavir discontinued prior to the interview so did not meet the inclusion/exclusion criteria and were not considered eligible and were excluded from the analyses. Thus there were three hundred ninety-nine subjects who were considered eligible and included in the analyses. Of these eligible subjects, ten subjects did not complete the interview.

Subject Demographics

In Table 9, it can be seen that the study included mostly male subjects in both the eligible and low literate subjects. For both the eligible subjects and the low literate subjects, very few of the subjects were 65 years of age or older. The most common race for both the eligible and low literate subjects was Caucasian/White. For eligible subjects, 57.6% of the subjects experienced nasal allergies within the last year. For low literate subjects, a lower percentage (39.1%) of the subjects experienced nasal allergies within the last year. Of the eligible subjects who suffered from indoor or outdoor allergies within the last year, 17.4% had used fluticasone propionate within that time. For the low literate subjects who suffered indoor or outdoor allergies during the last year, 22.2% had used fluticasone propionate during that time.

Table 9: Demographics for Eligible Subjects (RH01442)

	Total Eligible (N=309)		Low Literate (N=92)	
	n	%	n	%
Gender				
Male	339	85	70	76.1
Female	60	15	22	23.9
Age				
18 to 24	3	0.8	1	1.1
25 to 34	24	6	1	1.1
35 to 44	81	20.3	28	30.4
45 to 54	181	45.4	42	45.7
55 to 64	94	23.6	18	19.6
65 or older	16	4	2	2.2
Race				
Caucasian/White	273	68.4	45	48.9
African American/Black	79	19.8	38	41.3
Native American	6	1.5	0	0
Asian or Pacific Islander	5	1.3	1	1.1
Other	36	9	8	8.7
Hispanic Origin				
Yes	133	33.3	41	44.6

No	266	66.7	51	55.4
Education				
Less than High School	30	7.5	16	17.4
Completed High School	109	27.3	38	41.3
Some College/Technical School	97	24.3	18	19.6
Graduated College/Technical School or More	163	40.9	20	21.7
Income				
$0 to $14,999	220	55.1	74	80.4
$15,000 to $24,999	54	13.5	9	9.8
$25,000 to $34,999	29	7.3	3	3.3
$35,000 to $44,999	22	5.5	1	1.1
$45,000 to $64,999	16	4	0	0
$65,000 to $74,999	6	1.5	0	0
$75,000 or more	43	10.8	3	3.3
Refused	9	2.3	2	2.2
History (1yr) of Nasal Allergies				
Yes	230	57.6	36	39.1
No	169	42.4	56	60.9
Previous (1yr) FP Use*				
Yes	40	17.4	8	22.2
No	190	82.6	28	77.8

*Fluticasone propionate (FP) use in the past 12 months was assessed among all subjects reporting a recent history of indoor or outdoor nasal allergies.
Source: Study report, Table 2

3.2.3.4 Results and Conclusions (RH01442)

The correct self-selection rate in all eligible subjects was 43.6% (174/399) with a corresponding 2-sided 95% exact CI of (38.7%, 48.6%). This rate fell far below the performance threshold of 95%. As can be seen in Table 10, this low correct self-selection rate occurred regardless of literacy level, history of nasal allergies, and previous use of fluticasone propionate.

Table 10: Correct Self-selection rates (RH01442)

Population	**Correct Self-Selection Rate† % (n/N)**	**95% CI***
All eligible subjects	43.6 (174/399)	(38.7, 48.6)
Literacy level		
Low-literate	39.1 (36/92)	(29.1, 49.9)
Literate	45.0 (138/307	(39.3, 50.7)
History of nasal allergies		
Yes	43.0 (99/230)	(36.6, 49.7)
No	44.4 (75/169)	(36.8, 52.2)
Used FP previously		
Yes	37.5 (15/40)	(22.7, 54.2)

No	44.2 (84/190)	(37.0, 51.6)

† Correct self-selection rates is computed as #subjects with final code of "Correct" divided by the #eligible subjects
* 2-sided 95% exact CI
FP: fluticasone propionate
Source: Study report, Table 3

Upon further investigation, the Applicant tried to determine the reason why more than half of the subjects made an incorrect self-selection decision. For the 225 subjects who made an initial incorrect self-selection decision, the responses were categorized into the following categories:

- Comprehension (n=182, 80.9%),
- Label Prominence (n=24, 10.7%),
- Behavioral Override (n=16, 7.1%), and
- Other (n=3, 1.3%).

For the 225 subjects who made an incorrect self-selection decision, 101 (44.9%) said they were not prescribed ritonavir for the following reasons:

- Did not recognize their brand name drugs (Norvir, Kaletra) by generic name ritonavir (n=53, 23.6%), e.g. "I am taking Norvir not ritonavir",
- Was not prescribed ritonavir (n=39, 17.3%), e.g. "I'm not taking ritonavir any more that drug is old and out of date", or
- Did not know their drug by generic name ritonavir (n=9, 4%), e.g. "Did not know by that name".

The remaining 124 (55.1%) subjects who made an incorrect self-selection decision gave reasons not related to their prescribed HIV medication such as "I was not paying attention", "I would take because of allergies but I would ask my doctor first" and "The box says it's not okay. But my doctor says it is okay".

3.2.4 Human Factors Study (RH01801)

Minor revisions were made to the DFL label used in this study relative to the DFL used in the Label Comprehension and Self-Selection studies (RH01305, RH01318, and RH01442). These changes included, the term (b) (4) was replaced with "Quick Start Guide" in the "Directions" and the onset of action statement within "Other information" was modified slightly.

3.2.4.1 Study Design and Endpoints (RH01801)

The Applicant conducted a human factors study in adult subjects. Forty subjects were enrolled from a single site located in Raleigh, NC. The study focused on evaluating consumer understanding of the key "Use" instruction elements (how to prime, use and clean the product) via observation and interview and assessing behavioral aspects (how they execute against the key use instruction elements, including proper intranasal use vs. potential ocular use).

Subjects were enrolled and then given the test product, a placebo nasal spray that included packaging and labeling and were instructed to use the product as they would in their home. Subjects were observed for how they used the nasal spray in a simulated use environment (a home bathroom with a cabinet, sink, mirror, etc.). Participants did not receive any guidance or training prior to receiving the nasal spray. The investigator observed each subject as they worked

through the Initial Use session. After the Initial Use session was completed, the investigator then read instructions for the Two Week Later Use session. The primary intent of the Two Week Later Use session was to observe how subjects would clean the actuator after not having used the product for about "two weeks". They were then told to demonstrate what to do.

For the 'Initial Use' and 'Two Week Later Use' tasks, subjects were not prompted or guided and were free to use the product labeling (Drug Facts, Quick Start Guide) if they chose. Subjects who made an error during the "Initial Use" or "Two Week Later Use" sessions were directed to read the Quick Start Guide and asked to perform the tasks again. If the participant passed the task during the Repeat Failed Tasks section, it indicates that the Quick Start Guide provides sufficient information for the user to perform the tasks.

Primary objectives:
The primary objectives of the human factors study was to evaluate the use of the fluticasone propionate by a consumer in an OTC environment, specifically to understand the consumer's ability to clean and prime the nasal spray apparatus correctly, and to demonstrate that the consumer understands how to use the nasal spray, including the correct route of administration of the product (intranasal versus intraocular).

Primary endpoints
The endpoints for this study were the numbers of participants who had correctly used the product based on key criteria of the use instructions. Task performance was assessed for Initial Use session and the "Two week later use" session. Key items of interest for each session are listed below along with sub-steps involved with task:

- Initial Use
    - Dosed (nasal vs. ocular use)
    - Prime the pump
        - Cap removal, AND
        - Bottle shaking, AND
        - Pressing down & releasing the pump into air 6 times to see visible mist
    - Clean the actuator nozzle
        - Wipe the nozzle with a tissue OR
        - Rinse the nozzle with water
- Two Weeks Later Use
    - Dosed (nasal vs. ocular use)
    - Prime the pump
        - Cap removal, AND
        - Bottle shaking, AND
        - Pressing down & releasing the pump into air 6 times to see visible mist
    - Clean the actuator nozzle
        - Wipe the nozzle with a tissue OR
        - Rinse the nozzle with water
    - Clean the product after storing for at least a week
        - Remove the spray nozzle, AND
        - Rinse the spray nozzle under water, AND

- Replace the spray nozzle and correctly answering the action that they would take should the actuator be clogged.

3.2.4.2 Statistical Methodologies (RH01801)

The summative acceptance criterion was defined as 80% of participants able to successfully pass each task during the initial use session (Initial Use) and 80% of participants during the two week later use session (Two Week Later Use). Key items of interest included how the product was dosed (nasal vs. ocular use), priming of the pump and cleaning of the actuator nozzle as these attributes are essential for the safe and effective use of the product.

The steps were categorized as "Pass" or "Fail" as follows:

- Pass: successfully completed the task via the defined actions
- Fail: did not PASS / successfully complete the task.

At the completion of Initial Use and Two Week Later Use sessions, a single Repeat Failed Tasks session was conducted with subjects who failed any of the tasks. The investigator asked probing questions to better understand the reason why the participant was not successful in completing each task marked as a Fail.

During the Repeat Failed Tasks part of the session, the participant's actions to complete the steps based on reading the appropriate section of the Quick Start Guide were assessed as follows:

- Pass using instructions: completed the task using instructions
- Fail using instructions: did not complete the task after being referred to instructions

Pass/Fail criteria are listed below for each task:

Table 11: Pass/Fail criteria (RH01801)

Start of Task				Repeat?
Do they read the documentation?	Yes	No		Yes / No
Priming				Repeat?
Do they prime (mist appears)?	Yes	No		Yes / No
		Did they		
		Remove the cap?	Yes / No	
		Shake the bottle?	Yes / No	
		Press down & release pump into air 6 times to see mist?	Yes / No	
Blow Nose				Repeat?
Did they blow their nose?	Yes	No		Yes / No
Aim in nose?				Repeat?
Did they aim spray into one nostril AND hold other nostril with finger at the same time?	Yes	No		Yes / No
Did they aim spray into one nostril AND hold other nostril with finger at the same time?	Yes	No		Yes / No

Wipe Applicator			Repeat?
Did they wipe the nasal applicator with a clean tissue — If user cleaned via wiping OR rinsing via water — it would be considered a PASS	Yes	No	Yes / No
Cleaning			**Repeat?**
Remove the spray nozzle?	Yes	No	Yes / No
Soak the spray nozzle?	Yes	No	Yes / No
Rinse spray under water?	Yes	No	Yes / No
Replace the spray nozzle?	Yes	No	Yes / No
Ocular Use			**Repeat?**
Did the participant spray the product in the eyes?	Yes	No	Yes / No

Source: Study report, Appendix 2

During the post-task interview, subjects were asked if they had any difficulty using the nasal spray and whether anything was confusing. They were also asked for suggestions that would make the instructions better.

The Applicant provided the following justification for the 80% success criterion in a response to the Agency's 6/16/2014 Information Request:

> Typical acceptance criterion applied to usability testing ranges from 80 - 95% (ANSI/AMMI HFE75:2009 Human Factors Engineering – Design of Medical Devices, 2010). For studies RH01801 and RH01929, the lower criterion of 80% was chosen based on the low safety risk associated with failing to perform any of the key steps related to how the product was dosed (nasal vs. ocular use), priming of the pump and cleaning of the actuator nozzle and a reasonable estimate of the frequency at which these user errors might occur.
>
> For example, a failure to dose properly was considered to be of medium risk if sprayed directly in the eye. Although this action may cause user discomfort, it is not considered to be a significant safety concern as no safety issues have been identified as a result of post marketing reports of product misuse (spraying in eyes). Moreover, it is unlikely the consumer would repeat the action. Incomplete priming could result in suboptimal dosing which could affect efficacy, but would not be considered a safety concern. Similarly, improper cleaning could result in a clogged nozzle, which could subsequently lead to suboptimal dosing. However, this is not a safety concern and would most likely lead to a consumer re-reading the instructions to understand how to remove the clog or not using the product. Therefore, the clinical consequences associated with these identified failures were considered to be low in terms of overall product safety and any consequent risks appropriately managed through labeling.
>
> The report purposely provided actual pass/fail rates to assure full transparency of success and failure rates for each step during route of administration (intranasal versus intraocular), priming, and cleaning of the device. Although an 80% criterion was applied

to quantitatively assess the number of users completing identified tasks, the most important facet of human factors testing is the analysis of use errors and failures as a means to assess risk to the user and determine the need for risk mitigation. No matter what the acceptance criterion is set to and what level of pass/fail performance is observed, conducting a risk analysis for each and every usage error observed during the summative test is a priority.

As such, a thorough risk analysis was conducted on each and every usage error observed during testing - in several cases leading to enhancements to product labeling. The findings of these analyses are included within the study reports and support that the proposed Quick Start Guide provides sufficient information to enable consumers to use the product correctly.

Sample size
The Applicant stated that the sample size is based upon the ANSI/AMMI HFE75:2009 Human Factors Engineering – Design of Medical Devices guidance which describes that the standard number of participants for summative validation testing can be conducted with 15 -20 participants per user group.

3.2.4.3 Patient Disposition, Demographic and Baseline Characteristics (RH01801)

The demographic characteristics of the study are listed in Table 12. Subjects were evenly divided by gender. Very few of the subjects were 65 years of age or older [5% (2/40)]. The most common race was Caucasian/White [83% (33/83)]. Slightly more than half of subjects reported ocular symptoms [57.5% (23/40)]. More than half of the subjects reported nasal spray experience [60.0% (24/40)].

Table 12: Demographic Characteristics (RH01801)

	Total (N=40)	
	n	%
Gender		
Male	20	50
Female	20	50
Age		
18 to 24	6	15
25 to 34	8	20
35 to 44	8	20
45 to 54	10	25
55 to 64	6	15
65 or older	2	5
Race		
Caucasian/White	33	83
African American/Black	2	5
Asian or Pacific Islander	2	5
Other	3	8
Self-reported ocular symptoms		
Reported ocular symptoms	23	57.5
No reported ocular symptoms	17	42.5

Nasal spray experience		
Reported spray experience	24	60
No reported spray experience	16	40

Source: Study report, Tables 2 and 3

3.2.4.4 Results and Conclusions (RH01801)

The results for the key tasks are presented in Table 13. All of the subjects correctly understood the correct route of administration for the product (nasal vs. intraocular) at both the "Initial Use" session (40/40) and the "Two Weeks Later" session (40/40). The analyses for both sessions met the success criterion as the 2-sided exact 95% LCB of 91.2% for both sessions exceeded the 80% success criterion.

The performance of subjects for the prime task fell far below the pass threshold for both Initial Use [67.5% (27/40), 95% exact LCB = 50.9%] and Two Weeks Later [46.2% (18/39), 95% exact LCB=30.1%]. The reasons that subjects gave for skipping Priming included:

- Perceived waste of product,
- Not considered a necessary step,
- Something to do if clogged, and
- Not normal use.

Performance on the cleaning task that focused on wiping or rinsing the spray nozzle after each use also fell far below the 80% pass threshold for both the Initial Use session [52.5% (21/40), 95% exact LCB = 36.1%] and the Two Week Later Use session [46.2% (18/39), 95% exact LCB = 30.1%]. The reasons that subjects gave for not performing cleaning and wiping activities included:

- Single user for product (self),
- Would not occur to them, and not necessary for such a product.

The errors that were observed during this task were:

- Incorrectly put the bottle under the faucet and not the spray nozzle,
- Used an alcohol wipe to clean rather than tissue, and
- The subject did not see the instructions on the back of the Quick Start Guide.

The performance of subjects for the task of cleaning after storing the product for at least one week also fell far below the 80% threshold [11.4% (4/35), 95% exact LCB = 3.2%].

For the low performance in the cleaning tasks, the Applicant asserts that "While these errors could lead to poor hygiene and problems with clogging, none are considered a safety risk."

Table 13: Key Item Pass/Fail Percent for Initial and Two-week Later Use (RH01801)

Task	Initial Use		Two Weeks Later	
	% (n/N)	LB* (%)	% (n/N)	LB* (%)
Prime	67.5 (27/40)	50.9	46.2 (18/39)	30.1
Clean (Wipe/Rinse after each use)	52.5 (21/40)	36.1	46.2 (18/39)	30.1
Route of administration	100.0 (40/40)	91.2	100.0 (40/40)	91.2

Clean (after storing for at least 1 week)	NA	NA	11.4 (4/35)	3.2

* LB: 2-sided 95% exact lower confidence bound
Source: Study report, Table 4

If subjects failed a task during either the "Initial Use" or "Two Weeks Later Use" session, they were directed to review the Quick Start Guide. The results for the task performance after subjects reviewed the Quick Start Guide are presented in Table 14. Performance for both the prime [95.5% (21/22), 95% exact LCB = 77.2%] and wipe/rinse [92.9% (26/28), 95% exact LCB = 76.5%] tasks increased relative to the "Initial Use" and "Two Weeks Later" session but fell slightly below the 80% success criterion.

The performance for the cleaning after storing for at least one week task still fell far below the 80% threshold [63.6% (21/33), 95% exact LCB = 45.1%]. Performance on the individual sub-steps also all fell below the 80% threshold with the subjects having the most difficulty with the sub-step that asked subjects what they would if the nozzle became clogged.

Table 14: Repeated Key Items Pass/Fail Percent (RH01801)

Task	Correct Use % (n/N))	LB*
Prime	95.5 (21/22)	77.2
Wipe/Rinse	92.9 (26/28)	76.5
Clean (after storing at least 1 week)	63.6 (21/33)	45.1
Remove spray nozzle	87.9 (29/33)	71.8
Rinse spray nozzle	84.8 (28/33)	68.1
Replace spray nozzle	84.8 (28/33)	68.1
Soak nozzle in warm water if clogged	63.6 (21/33)	45.1

* LB: 2-sided 95% exact lower confidence bound
Source: Modified from Study report, Table 5

Subjects did not read the Quick Start Guide
During the Initial Use session, 13 out of 40 (33%) subjects failed to view the Quick Start Guide. During the Two Weeks Later Use session, 27 out of 39 (69%) subjects failed to view the Guide. To address this problem, the Applicant increased the prominence and accessibility of the Quick Start Guide. Also, to ensure that the Quick Start Guide is turned over, an arrow or statement was included to indicate that information continues on the back pages.

Inadvertent discharge of the product into the face
During the cleaning step twenty one subjects accidentally discharged the product when replacing the spray nozzle on the bottle. When questioned, all but one of these participants indicated that the discharge was a mistake or a surprise. In one instance, the one discharge went into the subject's face. The root cause for the inadvertent discharge toward the face was the position of the bottle during nozzle replacement after cleaning. In some instances the subject was holding the bottle with the spray nozzle pointed toward the face. In other instances the bottle was positioned on the counter and the participant was leaning over the bottle while replacing the

nozzle. The inadvertent spray always happened when subjects were replacing the spray nozzle after cleaning.

3.2.5 Human Factors Study in a Low Literacy Population (RH01929)

This second human factors study in a low literacy population was conducted because the Agency recommended that the Applicant conduct a human factors study in at least 15 low literate adults (May 2013 advice letter). Literacy was not prospectively measured in Study RH01801 so the Applicant conducted the second human factors study (Study RH01929). This study was conducted at a single site in Tarboro, NC.

3.2.5.1 Study Design and Endpoints (RH01929)

Study design, objectives, and endpoints are the same as the first human factors study conducted in a general population (RH01801). Please see Section 3.2.4.1 for details.

3.2.5.2 Statistical Methodologies (RH01929)

The statistical methods were the same as the first human factors study conducted in a general population. Please see Section 3.2.4.2 for details.

This study used the same 80% criterion pass rate as the first human factors study (RH01801). The criterion for this pass rate is included Section 3.2.4.2.

Sample size
The sample size of 15 subjects was consistent with the Agency recommendation that the human factors study be conducted in at least 15 low literate adults (May 2013 advice letter).

3.2.5.3 Patient Disposition, Demographic and Baseline Characteristics (RH01929)

The demographic characteristics of the study are listed in Table 15. There were several more females enrolled (n=9) than males (n=6). A minority of the subjects were 65 years of age or older [20% (3/15)]. The most common race was African American/Black [73% (11/15)]. One third (5/15) of the subjects reported ocular symptoms. More than half of the subjects reported nasal spray experience [53% (8/15)].

Table 15: Demographic Characteristics (RH01929)

	Total (N=15)	
	n	%
Gender		
Female	9	60
Male	6	40
Age		
19 to 24	3	20
25 to 34	1	7
35 to 44	2	13
45 to 54	2	13
55 to 64	4	27
65 or older	3	20
Race		

Caucasian/White	4	27
African American/Black	11	73
Self-reported ocular symptoms		
Reported ocular symptoms	5	33
No reported ocular symptoms	10	67
Nasal spray experience		
Reported spray experience	8	53
No reported spray experience	7	47

Source: Study report, Tables 2 and 3

3.2.5.4 Results and Conclusions (RH01929)

The results for the key tasks are presented in Table 16. Subjects correctly understood that the product was for nasal use and not for ocular use. No subject intentionally sprayed the product into their eyes at either the Initial Use session or the Two Week Later session.

The performance of subjects for the prime task fell far below the pass threshold for both Initial Use [40.0% (6/15), 95% exact LCB = 16.3%] and Two Week Later Use [33.3% (5/15), 95% exact LCB=11.8%] sessions. The reasons that subjects gave for skipping the Priming included:

- Not something they currently do with their nasal spray,
- Misinterpreted instructions in the Quick Start Guide – Participant indicated they skipped Prime because it said to go to Step 3 unless they just cleaned the nozzle, and
- Not considered a necessary step, and stated they forgot.

Performance on the cleaning task that focused on wiping or rinsing the spray nozzle after each use also fell far below the 80% success criterion for both the Initial Use [33.3% (5/15), 95% exact LCB = 11.8%] and Two Week Later [40.0% (6/15), 95% exact LCB = 16.3%] sessions. The reasons that subjects gave for not performing cleaning and wiping activities included:

- Single user for product (self),
- Would not occur to them,
- Not necessary for such a product, and
- Skipped over it in the Quick Start Guide, and
- Stated they forgot.

The errors that were observed during this task included:

- Participant did not see the instructions on the back of the Quick Start Guide,
- Participant misinterpreted the text in the Quick Start Guide thinking the “spray nozzle” was actually referencing the green cap which resulted in cleaning with the spray nozzle remaining on,
- Participant only completed steps 1 and 3, skipping step 2 and the instructions on soaking if clogged, and
- Incorrectly stating a method for unclogging the sprayer (e.g., using a pin).

The Applicant asserts that “While these errors could lead to poor hygiene and problems with clogging, none are considered a safety risk.”

No subject (0/15) correctly completed the task of cleaning after storing the product for at least one week.

Table 16: Key Item Pass/Fail Percent for Initial and Two-week Later Use (RH01929)

Task	Initial Use		Two Weeks Later	
	% (n/N)	LB*	% (n/N)	LB*
Prime	40.0 (6/15)	16.3	33.3 (5/15)	11.8
Clean (Wipe/Rinse after each use)	33.3 (5/15)	11.8	40.0 (6/15)	16.3
Route of administration	100.0 (15/15)	78.2	100.0 (15/15)	78.2
Clean (after storing for at least 1 week)	NA		0.0 (0/15)	0.0

*LB: 2-sided 95% exact lower confidence bound
Source: modified from Study report, Table 4

If subjects failed a task during either the "Initial Use" or "Two Weeks Later Use" session, they were directed to review the Quick Start Guide. The results for the task performance after subjects reviewed the Quick Start Guide are presented in Table 17. All subjects who repeated the Prime (10/10) and Wipe or Rinse (11/11) tasks correctly completed the tasks. Performance on the clean after storing at least one week task increased substantially after the subjects reviewed the Quick Start Guide but still only slightly more than half of the subjects [7/13 (54.0%)] correctly completed all of the cleaning sub-steps with subjects having the most difficulty with the sub-step that asked subjects what they would do if the nozzle became clogged.

Table 17: Repeated Key Elements Pass/Fail Percent (RH01929)

Task	Correct Use	
	% (n/N)	LB*
Prime	100.0 (10/10)	69.2
Wipe/Rinse	100.0 (11/11)	71.5
Clean (after storing at least 1 week)	53.8 (7/13)	25.1
Remove spray nozzle	92.3 (12/13)	64.0
Rinse spray nozzle	84.6 (11/13)	54.6
Replace spray nozzle	92.3 (12/13)	64.0
Soak nozzle in warm water if clogged	53.8 (7/13)	25.1

Source: Modified from Study report, Table 5

During the cleaning step, eight subjects discharged product accidentally when replacing the spray nozzle on the bottle. Unlike the previous study (RH01801) where inadvertent discharges occasionally were directed at the face, no inadvertent discharges towards the face occurred in this study. In fact, ten of the thirteen subjects who repeated the cleaning step were observed purposefully pointing the bottle away from the face while replacing the sprayer nozzle.

3.3 Evaluation of Safety

Subjects in the consumer behavior studies were not exposed to treatment; therefore, the evaluation of safety is not applicable for these studies.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

All consumer behavior studies were conducted in the United States only so subgroup analyses by geographic region were not conducted.

There were no meaningful subgroups differences for gender, age, and race observed in Studies RH1305, RH1318, and RH1442.

Subgroup analyses were not conducted in the human factors studies (RH01801 and RH01929) because the small number of subjects in these studies precluded meaningful subgroup analyses.

4.2 Other Special/Subgroup Populations

The other important subgroup analyses have been discussed previously in Section 3.2.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

Self-selection in HIV patients taking ritonavir

In the targeted self-selection study (RH01442) of adult HIV patients taking ritonavir, the correct self-selection rate was 43.6% (174/399) with a corresponding 2-sided 95% exact CI of (38.7%, 48.6%). This correct self-selection rate fell far below the 95% performance threshold. The low correct self-selection rate occurred regardless of literacy level, history of nasal allergies, and previous use of fluticasone propionate. The low rate of correct selection is concerning because concomitant use of fluticasone propionate and ritonavir could place consumers at increased risk for systemic glucocorticoid side-effects.

To aid in improving consumer understanding that fluticasone propionate and ritonavir should not be co-administered, (b) (4)

(b) (4)

(b) (4)

Poor performance in the cleaning task for the human factors studies

Performance of subjects for the task of cleaning after storing the product for at least one week in the general population fell far below the 80% threshold [11.4% (4/35), 95% exact LCB = 3.2%]. Similar poor results were seen in the low literate population where no subjects correctly completed the tasks involved with cleaning after storing the device for at least one week.

Many subjects did not view the Quick Start Guide during both the "Initial Use" session [33% (13/40)] and "Two Week Later Use" session [69% (27/30)]. They posited that if more subjects reviewed the Quick Start Guide, this would increase correct performance of the tasks as evidenced by the higher rates observed in the Repeat session. To ensure that consumers would be more likely to review the Quick Start Guide, the Applicant increased the prominence and accessibility of the Quick Start Guide. Also, to ensure that the Quick Start Guide was turned over, an arrow or statement was included to indicate that information continues on the back pages. It should be noted that, while the Applicant expected that these changes would increase consumer review of the Quick Start Guide and improve task performance, the modified labelling was not tested.

In the general population, during the Repeat session for subjects who failed the clean after storing at least one week during either the "Initial Use" or "Two Weeks Later Use" sessions, performance increased but still fell far below the 80% success criterion [63.6% (21/33), 95% exact LCB = 45.1%]. In the low literate population, during the Repeat session, the performance for this task also fell far below the 80% success criterion [53.8% (7/13), 95% exact LCB = 25.1%]. In looking at the performance on individual sub- steps for this task, subjects in both the general and low literate population by far had the most difficulty with the sub-step that asked subjects what they would do if the nozzle became clogged. The results for "Repeat" session suggest that the changes to the packaging to increase consumer review of the Quick Start Guide will likely increase the rate of proper cleaning but does not provide confidence that it would meet the 80% success criterion. The Applicant stated that "While these errors could lead to poor hygiene and problems with clogging, none are considered a safety risk."

Improper cleaning that leads to clogging will likely result in decreased efficacy and possibly in fewer purchases. The determination of whether errors in cleaning would be considered a safety risk is left to the clinical reviewer.

5.2 Collective Evidence

Comprehension of warnings and instructions

The pilot label comprehension study (RH01305) and the targeted label comprehension study (RH01318) were conducted sequentially with the label used in the targeted label comprehension study modified based on the results of the pilot label comprehension study. The modifications are detailed in Section 3.2.2.

In both the pilot label comprehension study and targeted label comprehension study, most subjects in the general population as well as the low literate population correctly understood the dosing directions and comprehension for these items met the 90% performance threshold. Almost all of the subjects in the pilot comprehension study in both the general population (105/106) and the low literate population (30/31) correctly understood the uses of the product. Similarly, for the question, "Ask a doctor or pharmacist before use if you are taking ritonavir", almost all of the subjects in the pilot label comprehension study for both the general population (105/106) and low literate population (30/31) correctly comprehended this objective.

For the "Do not use to treat asthma" objective in the pilot label comprehension study, most general population subjects understood this objective [96.2% (102/106), 95% exact LCB=90.6%]. Comprehension for the low literate population was also relatively high [93.5% (29/31), 95% exact LCB=78.6%].

In the pilot label comprehension study, comprehension of the "Use this product only once daily" objective fell slightly below the 90% performance threshold for the general population [92.5% (98/106), 95% exact LCB=85.7%] and was lower for the low literate population [80.6% (25/31), 95% exact LCB=62.5%].

In the pilot label comprehension study, comprehension of the "Stop use and ask a doctor if your symptoms do not get better within 7 days of starting use" warning was not that well understood (General population: 79.2%, 95% exact lower confidence bound (LCB) = 70.3%; low literate population: 54.8%, 95% exact LCB = 36.0%). However, in the subsequent targeted label comprehension study, subjects had much higher comprehension of the warning (General population: 91.0%, 95% exact lower confidence bound (LCB) = 88.2%; low literate population: 93.5%, 95% exact LCB =88.3%).

In both the pilot and targeted label comprehension studies, comprehension of the "Stop use and ask a doctor if you get new symptoms such as severe facial pain" warning was high (Pilot: 95.3%, 95% exact LCB = 89.3%; Targeted: 95.1%, 95% exact LCB=92.8%). Comprehension rates for this warning were slightly lower in the low literate population (Pilot: 80.6%, 95% exact LCB =62.5 %; Targeted: 85.0%, 95% exact LCB=78.3%).

Self-selection in HIV patients taking ritonavir

This was studied in Study RH01442. (b) (4)

Performance in the human factors studies

All of the subjects in the general (RH01801) and low literate (RH01929) populations correctly understood the correct route of administration for the product (nasal vs. intraocular) at both the

"Initial Use" session (40/40 for the general population and 15/15 for the low literate population) and the "Two Weeks Later Use" session (40/40 and 15/15).

For the human factor study in the general population (RH01801), subjects' performance for the priming task fell far below the 80% success criterion for both Initial Use [67.5% (27/40), 95% exact LCB = 50.9%] and Two Weeks Later [46.2% (18/39), 95% exact LCB=30.1%] sessions. In the low literate population (RH01929), performance also fell far below the 80% success for both Initial Use [40.0% (6/15), 95% exact LCB=16.3%] and "Two Weeks Later Use" [33.5% (5/15), 95% exact LCB=11.8%] sessions.

In the general population, performance on the cleaning task that focused on wiping or rising the spray nozzle after each use fell far below the 80% success criterion for both the Initial Use [52.5% (21/40), 95% exact LCB = 36.1%] and Two Week Later Use [46.2% (18/39), 95% exact LCB = 30.1%] sessions. Performance in the low literate population also fell far below the 80% success criterion for both the Initial Use [33.3% (5/15), 95% exact LCB=11.8%] and Two Weeks Later Use [40.0% (6/15), 95% exact LCB=16.3%] sessions.

In the general population, the results for subjects who failed a task during either the "Initial Use" or "Two Weeks Later" session and were directed to review the Quick Start Guide (the "Repeat Failed Tasks" session) were much higher for both the "Prime" [95.5% (21/22), 95% exact LCB = 77.2%] and "Wipe/Rinse [92.9% (26/28), 95% exact LCB = 76.5%] tasks with both tasks falling slightly below the 80% success criterion. For subjects who initially failed a task in either the "Initial Use" or "Two Week Later" sessions in the low literate population, all of the subjects correctly performed the Prime (10/10) and Wipe/Rinse (11/11) tasks . The Applicant viewed these analyses as an assessment of the Quick Start Guide's ability to aid consumer understanding of correct product use.

Because many subjects did not initially review the Quick Start Guide and it increased correct performance for many of the tasks, the Applicant proposed changes to the packaging to increase review of the Quick Start Guide. These changes are listed in Section 5.1.

Poor performance in the cleaning task for the human factors studies
The poor performance of subjects for the task of cleaning after storing the product for at least one week fell far below the 80% success criterion for the "Initial Use" and "Two Week Later" sessions. During the "Repeat" session for subjects who made errors during either the "Initial Use" or "Two Weeks Later Use" session, performance improved but still fell far below the 80% success criterion with subjects having the most difficulty with the sub-step that asked subjects what they would do if the nozzle became clogged. The results for "Repeat" session suggest that the changes to the packaging to increase consumer review of the Quick Start Guide will likely increase the rate of proper cleaning but does not provide confidence that it would meet the 80% success criterion.

5.3 Conclusions and Recommendations
Overall subjects in the general and low literate populations generally understood the warnings and directions in the label. There were issues with subjects' understanding that they should not take fluticasone propionate and ritonavir together. (b) (4)

(b) (4)

Subjects had some difficulty in correctly cleaning the product after storing it for at least one week. Most errors occurred in the sub-step where subjects stated how they would unclog the device if a clog occurred. The determination of whether errors in cleaning resulting in clogged devices would be considered a safety risk is left to the clinical reviewer.

9 Page(s) of Draft Labeling have been Withheld in Full as b4 (CCI/TS) immediately following this page

---

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

/s/

SCOTT S KOMO
06/26/2014

KAREN M HIGGINS
06/26/2014
I concur.

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 Serial Number:	205434/0000
Drug Name:	Flonase (fluticasone propionate aqueous nasal spray)
Indication(s):	Treatment of Ocular Symptoms of Allergic Rhinitis
Applicant:	GlaxoSmithKline (GSK)
Date(s):	Received: September 23, 2013 PDUFA: July 23, 2014
Review Priority:	Standard
Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	David Hoberman, PhD
Concurring Reviewers:	David Petullo, MS
Medical Division:	Division of Pulmonary and Rheumatoid Arthritis Products
Clinical Team:	Medical Officer: Stacy Chen, MD Medical Team Leader: Susan Limb, MD
Project Manager:	Jung E Lee, RPh

Keywords: clinical studies, supplemental NDA review, double-blind

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1. EXECUTIVE SUMMARY

As part of the over-the-counter switch of Flonase GSK has proposed an OTC label that includes relief of ocular symptoms associated with allergic rhinitis. However the efficacy of Flonase in relieving ocular symptoms has not been established. Therefore the applicant conducted three studies to evaluate relief of ocular symptoms associated with allergic rhinitis, trials 30033, 30034, and RH01619.

Trial 30033 demonstrated a statistically significant difference between Flonase 200 mcg QD (FP200QD) and placebo with respect to the primary endpoint, total ocular symptom score (TOSS). However, in Trial 30034 the results were marginal, p-value=0.055. This was likely due to a considerably greater placebo response than observed in Trial 30033. However, the average TOSS of the FP200QD group over 24 weeks in both studies was similar. There was no substantial evidence of treatment by gender interaction or treatment by age interaction.

The results from Trial RH01619 were somewhat anomalous despite the statistically significant difference between FP200QD and placebo in this two week study.

- Examination of the empirical distribution functions (Figure 7) does not show a clean location shift. Treatment differences appear only among those who improved from baseline. Thus the p-value of 0.0024 may detect a signal of efficacy but the interpretation is in doubt.
- There was no relationship between average performance of treatment arm and different pollen counts at the six sites, Figure.
- On average, males derived no benefit from FP200QD when considering the primary endpoint, TOSS.

Taken together, these three findings cast doubt on the utility of this trial to support the efficacy of Flonase for the relief of ocular symptoms due to allergic rhinitis.

2. INTRODUCTION

2.1 Overview

The efficacy of Flonase for the relief of nasal allergy symptoms has been firmly established however the relief of ocular symptoms has not been established. The applicant submitted three studies to support the efficacy of Flonase in relieving ocular symptoms associated with allergic rhinitis, Studies 30033, 30034, and RH01619. These three studies are the focus of my review.

2.2 Data Sources

All data was supplied electronically by the Applicant as SAS transport files and can be found at the following location in the CDER electronic document room (EDR):

\\Cdsesub1\EVSPROD\NDA205434\0000\m5\datasets

3. STATISTICAL EVALUATION

Studies 30033 and 30034 were randomized double-blind, placebo- and active-controlled, 4-week trials. Study RH01619 was a 2-week trial that did not contain an active-control. These studies are evaluated separately under Sections 3.2.1, 3.2.2, and 3.2.3.

3.1 Data and Analysis Quality

The electronic data submitted by the applicant for the three studies was of sufficient quality to allow a thorough review of the data. I was able to derive the primary and secondary endpoints for each study. The statistical analyses of my derived endpoints were in agreement with the Applicant's analyses.

3.2 Evaluation of Efficacy

My review focuses on the three studies submitted by the applicant to support the efficacy of Flonase in relieving ocular symptoms associated allergic rhinitis. Each study is discussed separately in section 3.2.1, 3.2.2, and 3.2.3.

3.2.1 Study 30033

Study Design and Endpoints

This was a randomized, double-blind, double-dummy, placebo- and active-controlled, parallel-group trial in subjects with seasonal allergic rhinitis that was conducted from March 2001 – July 2001. This trial was four weeks in duration. Male and female subjects that were 12 years of age, had a positive skin test, historical evidence of seasonal allergic rhinitis (for the past 2 years), and had a total ocular symptom score of 120 (out of 330) and a nasal congestion score of 50 (out of 100) were eligible for enrollment. After completing a screening period, subjects were randomly assigned to placebo, FP200QD, or loratadine. Subjects visited the clinic at screening and at Days 1, 14 and 28. The primary endpoint is the mean change from baseline in reflective subject-rated total ocular symptom scores (TOSS). TOSS is defined as the sum of the scores from three symptoms, itching, tearing, and redness and each symptom is scored using a 100 mm VAS. Secondary endpoints include mean change from baseline in reflective subject-rated scores for both individual ocular symptoms (over Days 1-28,1-7, 8-14, 15-21, and 22-28) and total ocular symptoms (Days 1-7, 8-14, 15-21, and 22-28).

The applicant estimated that a sample size of 152 subjects per treatment group would detect a treatment difference of 25.5 units between placebo and FP200QD in the change from baseline in TOSS. They assumed a standard deviation of 73.8.

Patient Disposition, Demographic and Baseline Characteristics

This study enrolled 471 subjects among 14 centers in the United States. Demographics and baseline characteristics for all randomized and treated patients are shown in Table 1.

Table 1. Demographics and baseline characteristics for Study 30033

Demography and Subject Characteristics

	Placebo (N=155)	FP200QD (N=158)	LOR10QD (N=158)	Total (N=471)
Age Group, n(%)				
Less than 18	38 (25%)	16 (10%)	29 (18%)	83 (18%)
18-64	113 (73%)	138 (87%)	128 (81%)	379 (80%)
65-74	4 (3%)	3 (2%)	1 (<1%)	8 (2%)
75 or more	0	1 (<1%)	0	1 (<1%)
Age (years)				
n	155	158	158	471
Mean(sd)	32.3(15.2)	35.5(13.9)	32.9(13.4)	33.6(14.2)
Median	31.0	34.0	33.0	33.0
Min-Max	12-71	12-76	12-69	12-76
Sex, n(%)				
Female	91 (59%)	98 (62%)	91 (58%)	280 (59%)
Male	64 (41%)	60 (38%)	67 (42%)	191 (41%)
Ethnic Origin, n(%)				
White	128 (83%)	140 (89%)	127 (80%)	395 (84%)
Black	20 (13%)	13 (8%)	16 (10%)	49 (10%)
Asian	2 (1%)	2 (1%)	5 (3%)	9 (2%)
American Hispanic	4 (3%)	3 (2%)	9 (6%)	16 (3%)
Other	1 (<1%)	0	1 (<1%)	2 (<1%)
Height (cm)				
n	155	158	158	471
Mean(sd)	168.5(10.5)	168.2(9.9)	168.4(10.2)	168.4(10.2)
Median	168.0	168.0	168.0	168.0
Min-Max	139-195	147-201	142-193	139-201
Weight (kg)				
n	155	158	158	471
Mean(sd)	76.5(21.0)	77.9(20.4)	77.0(21.5)	77.1(20.9)
Median	73.2	75.0	73.0	73.6
Min-Max	36-135	37-135	38-186	36-186

Source: Table 5 from applicant's Clinical Study Report

This study comprised of mostly Caucasian subjects with a mean age in years of approximately 33. There were slightly more female subjects than male subjects but was evenly distributed between treatment arms. There was no significant difference in baseline characteristics.

There was over 90% completion rate in this study. Reasons for discontinuation are shown in Table 2.

Table 2. Disposition of subjects in Study 30033

Subject Disposition

	Placebo (N=155)	FP200QD (N=158)	LOR10QD (N=158)	Total (N=471)
Completion status, n(%)				
Completed	141 (91%)	149 (94%)	137 (87%)	427 (91%)
Prematurely discontinued	14 (9%)	9 (6%)	21 (13%)	44 (9%)
Reason for Premature Discontinuation, n(%)				
Adverse event	3 (2%)	5 (3%)	6 (4%)	14 (3%)
Consent withdrawn	2 (1%)	0	0	2 (<1%)
Lost to follow up	0	1 (<1%)	4 (3%)	5 (1%)
Protocol violation	3 (2%)	0	2 (1%)	5 (1%)
Lack of efficacy	5 (3%)	0	6 (4%)	11 (2%)
Other	1 (<1%)	3 (2%)	3 (2%)	7 (1%)
Pregnancy During Study, n(%)	1 (1%)	0	0	1 (<1%)

Source: Table 2 from applicant's clinical study report

The over 90% completion rate, together with the large sample sizes in each group and the averaging over time, obviate the need for detailed considerations regarding missing data.

The duration in years the subjects suffered from seasonal allergic rhinitis is shown in Table 3.

Table 3. Duration of seasonal allergic rhinitis for subjects enrolled in Study 30033

Duration of Seasonal Allergic Rhinitis

	Placebo (N=155)	FP200QD (N=158)	LOR10QD (N=158)
Duration of seasonal allergic rhinitis, n(%)			
Less than 2 years	0	0	0
2 to 5 years	21 (14%)	24 (15%)	11 (7%)
6 to 10 years	36 (23%)	25 (16%)	39 (25%)
11 to 15 years	30 (19%)	21 (13%)	32 (20%)
16 to 20 years	18 (12%)	19 (12%)	14 (9%)
21 to 25 years	17 (11%)	16 (10%)	13 (8%)
26 or more years	33 (21%)	53 (34%)	49 (31%)
Duration of seasonal allergic rhinitis (years)			
n	155	158	158
Mean(sd)	17.4(12.6)	19.5(12.6)	19.0(12.2)
Median	13.0	17.0	15.0
Min-Max	2-62	2-51	2-52

Source: Table 6 from applicant's clinical study report

The duration of seasonal allergies was similar across treatment arms.

Statistical Methodologies

The analysis dataset defined by the applicant was all randomized subjects. The derived data for each subject consisted of the average TOSS over the course of each week and over the entire 4 weeks (28 days). The primary statistical endpoint was the average change from baseline of the TOSS over the entire 28 days. The primary analysis method was ANCOVA with baseline of the measure being analyzed and investigator.

Secondary endpoints included averages over the aforementioned 4 weekly epochs for each symptom separately, and the TOSS over the four separate weekly epochs. Results for subject-rated overall evaluation of response to treatment were compared between treatment groups using a Cochran-Mantel-Haenszel (CMH) adjusted for site. The hierarchical plan for testing multiple secondary endpoints is shown in Table 4.

Table 4. Sequential testing strategy for secondary endpoints in Study 30033

The primary efficacy measure, the mean change from baseline in TOSS compared averaged over Days 1-28 between FP200QD and placebo, will be assessed at a significance level of $\alpha = 0.05$. If this comparison is statistically significant, then the secondary efficacy measures will be assessed for statistical significance, based on the following hierarchical closed testing strategy and disclosure of results scheme.

If the comparison of FP200QD and placebo for the primary endpoint is statistically significant at $\alpha = 0.05$, then the secondary measures, mean change from baseline in the individual symptoms: itching, tearing, and redness over Days 1-28 for FP200QD compared with placebo, will each be assessed at $\alpha = 0.05$. In addition, if the comparison of FP200QD and placebo over Days 1-28 for a given secondary measures (i.e., itching, tearing, or redness) is statistically significant at $\alpha = 0.05$, then mean change from baseline for that symptom over Days 22-28, Days 15-21, Days 8-14, and Days 1-7, in that order, will be assessed at $\alpha = 0.05$. Statistical significance must be achieved at a $\alpha = 0.05$ for a given week in order to progress to the next week in the hierarchy. Weekly mean changes from baseline in TOSS will be assessed using the same hierarchical rules.

For the three overall secondary measures (I.e., overall mean changes for ocular itching, tearing, and redness), the following disclosure of results scheme will be employed. Reporting the results of one overall mean change requires reporting of all other results within the set (i.e., the remaining individual overall mean changes), regardless of the statistical significance associated with the test measure. Conversely, choosing not to report any one measure requires that none of the other measures within the set be reported.

There will be no multiplicity adjustment for the set of exploratory endpoints which includes the mean changes from baseline in nasal congestion, the end of treatment subject-rated overall evaluation of response to therapy, and the exploration of the treatment difference in the ocular symptom endpoints comparing FP200QD with LOR10QD and LOR10QD with placebo.

Source: Applicant's protocol, page 20

There was no imputation of missing data for either the primary or secondary analyses.

Results and Conclusions

The results for the comparison of TOSS over 4 weeks and at each week are shown in Table 5.

Table 5. Results from analyses of TOSS in Study 30033

Change from Baseline in Subject-Rated Total Ocular Symptom Scores

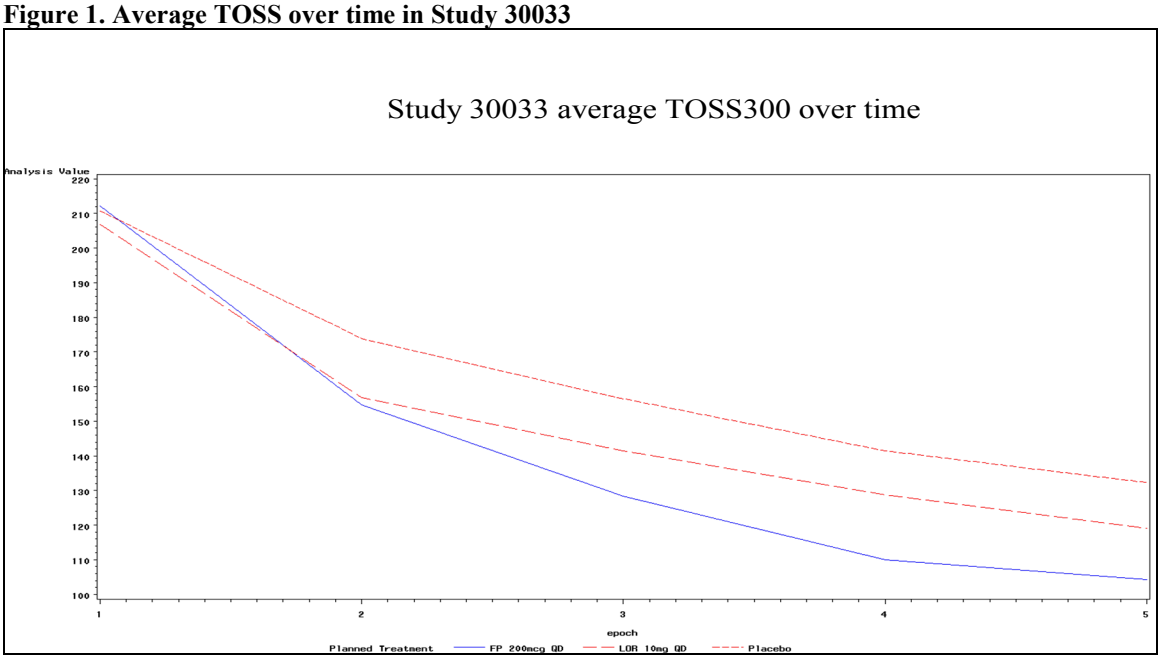
	Placebo (N=155)	FP200QD (N=158)	LOR10QD (N=158)	FP200QD vs Placebo[1]	LOR10QD vs Placebo[1]	FP200QD vs LOR10QD[1]
Baseline[2]						
n	155	158	158			
Mean adj(se)	207.9(3.8)	209.4(3.8)	204.3(3.8)			
Days 1-7						
n	155	156	155	0.003	0.036	0.357
Mean adj(se)	-39.1(5.0)	-59.7(4.9)	-53.4(5.0)			
Days 8-14						
n	153	155	153	<0.001	0.126	0.045
Mean adj(se)	-57.4(6.1)	-86.6(6.0)	-70.0(6.0)			
Days 15-21						
n	144	153	144	<0.001	0.213	0.011
Mean adj(se)	-72.6(6.3)	-105.1(6.1)	-83.4(6.3)			
Days 22-28						
n	142	150	138	0.001	0.256	0.038
Mean adj(se)	-82.8(6.4)	-111.2(6.2)	-92.9(6.5)			
Days 1-28						
n	155	157	155	<0.001	0.086	0.028
Mean adj(se)	-59.9(5.4)	-88.7(5.3)	-72.5(5.4)			

Note: Total Ocular Symptom Score is the sum of itching, tearing and redness scores.
Note: Values shown are actual values at baseline, and change from baseline for subsequent time periods.
Note: Subjects 8196, 8284, 8288, and 9007 did not provide any diary data during treatment.
[1] p-values are based on analysis of covariance on change from baseline, adjusting for investigator and baseline value, using linear contrasts for pairwise comparisons.
[2] Baseline is defined as the average over the 7 days immediately preceding Day 1 (i.e., treatment start date).

Source: Table 10 from applicant's clinical study report

Note that the p-value for the comparison of FP200QD versus placebo over the entire 4 weeks and the last 7 days of treatment (Days 22-28) is 0.001. The same pattern of results is observed for comparison each individual symptom, both averaging over days 1-28 and averaging over the last epoch: days 22-28. As an adjunct analysis, the reviewer found that a principle components analysis revealed that each symptom contributed equally to the variation in TOSS scores. The mean TOSS scores over time are shown in Figure 1. Epoch 1= days -6 to 0 (baseline), epoch 2=days 1-7, epoch 3 = days 8-14, epoch 4= days 15-21, and epoch 5= days 22-28. Note TOSS300 refers to TOSS. Note TOSS300 refers to TOSS.

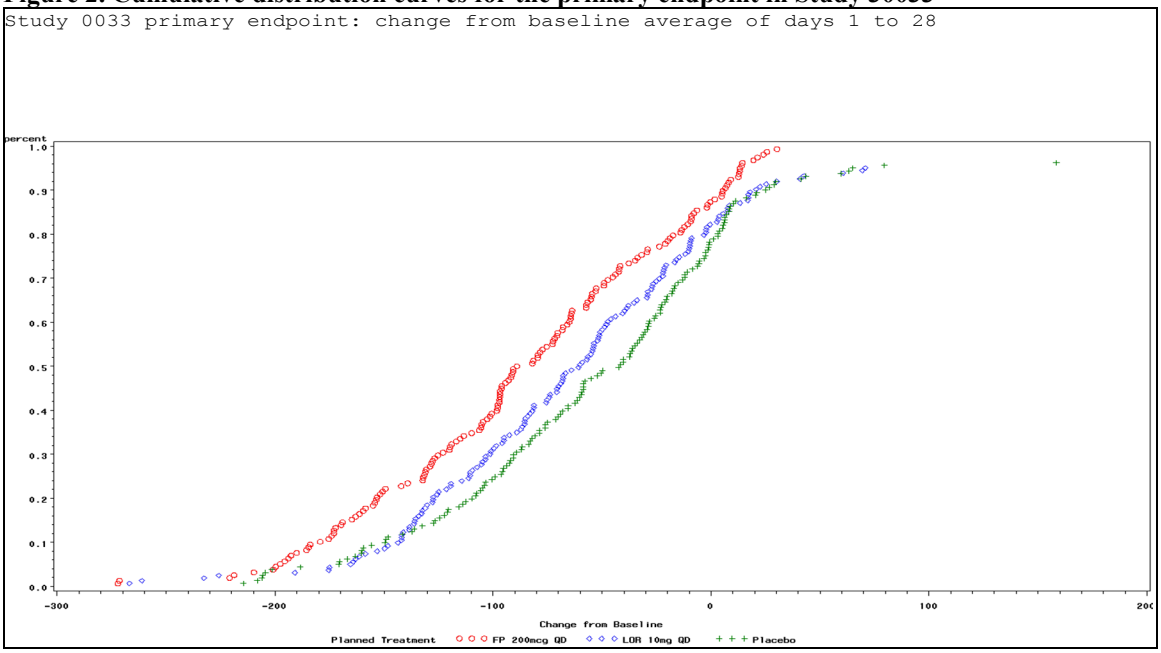
Figure 1. Average TOSS over time in Study 30033



Source: Reviewer

There was clear separation between FP200QD and placebo for all time periods examined. The cumulative distribution curves for the primary endpoint (TOSS averaged over days 1-28) are shown in Figure 2.

Figure 2. Cumulative distribution curves for the primary endpoint in Study 30033



Source: Reviewer

As observed with the change in TOSS, there was clear separation between placebo and FP200QD for all levels of response.

Results from the comparison of subject-rated overall evaluation of response to treatment are shown in Table 5.

Table 6. Results for the analysis of subject-rated overall evaluation of response to treatment in Study 30033

Subject-Rated Overall Evaluation of Response to Therapy

	Placebo (N=155)	FP200QD (N=158)	LOR10QD (N=158)	FP200QD vs Placebo[1]	LOR10QD vs Placebo[1]	FP200QD vs LOR10QD[1]
Evaluation of response, n(%)				<0.001	0.524	<0.001
n	155	157	154			
Significantly improved	12 (8%)	45 (29%)	20 (13%)			
Moderately improved	43 (28%)	35 (22%)	45 (29%)			
Mildly improved	45 (29%)	49 (31%)	33 (21%)			
No change	43 (28%)	17 (11%)	40 (26%)			
Mildly worse	7 (5%)	4 (3%)	7 (5%)			
Moderately worse	3 (2%)	5 (3%)	6 (4%)			
Significantly worse	2 (1%)	2 (1%)	3 (2%)			

Source: Table 15 from applicant's clinical study report

There was a significant difference between placebo and FP200QD, p-value < .001.

The p-value of .001 for the comparison of Flonase to placebo for the primary endpoint, TOSS, supports the efficacy of Flonase in relieving ocular symptoms associated with allergic rhinitis. This was supported by the secondary endpoint, subject-rated overall evaluation of response to treatment. Further, there was clear separation in the cumulative distribution of responders in favor of Flonase.

3.2.1. Study 30034

This study was conducted from March 2001 - June 2001.

Study Design and Endpoints

The design and analyses of this study was identical to Study 30033. See section 3.2.1

Patient Disposition, Demographic and Baseline Characteristics

This study randomized 482 subjects among 14 centers. Demographics and baseline characteristics for all randomized and treated patients are shown in Table 7.

Table 7. Demographics and baseline characteristics for subjects in Study 30034

Demography and Subject Characteristics

	Placebo (N=161)	FP200QD (N=158)	LOR10QD (N=163)	Total (N=482)
Age Group, n(%)				
Less than 18	33 (20%)	29 (18%)	22 (13%)	84 (17%)
18-64	122 (76%)	127 (80%)	134 (82%)	383 (79%)
65-74	5 (3%)	2 (1%)	5 (3%)	12 (2%)
75 or more	1 (<1%)	0	2 (1%)	3 (<1%)
Age (years)				
n	161	158	163	482
Mean(sd)	34.1(15.2)	35.4(14.1)	35.1(14.7)	34.9(14.7)
Median	33.0	37.0	34.0	35.0
Min-Max	12-79	12-69	12-75	12-79
Sex, n(%)				
Female	99 (61%)	106 (67%)	105 (64%)	310 (64%)
Male	62 (39%)	52 (33%)	58 (36%)	172 (36%)
Ethnic Origin, n(%)				
White	103 (64%)	103 (65%)	113 (69%)	319 (66%)
Black	29 (18%)	26 (16%)	26 (16%)	81 (17%)
Asian	2 (1%)	3 (2%)	4 (2%)	9 (2%)
American Hispanic	27 (17%)	25 (16%)	19 (12%)	71 (15%)
Other	0	1 (<1%)	1 (<1%)	2 (<1%)
Height (cm)				
n	161	157	163	481
Mean(sd)	168.1(10.3)	167.3(8.8)	167.4(8.8)	167.6(9.3)
Median	168.0	168.0	166.0	168.0
Min-Max	137-191	145-191	144-193	137-193
Weight (kg)				
n	161	156	163	480
Mean(sd)	78.7(21.0)	75.2(18.7)	77.0(16.9)	77.0(19.0)
Median	75.0	72.5	77.2	75.5
Min-Max	35-168	33-141	42-126	33-168

Source: Table 5 from applicant's clinical study report

The average age in years for all subjects was approximately 35 and as in Study 30033, there were slightly more female subjects than male subjects. Duration of seasonal allergic rhinitis for randomized subjects is shown in Table 8.

Table 8. Duration of seasonal allergic rhinitis in Study 30034

Duration of Seasonal Allergic Rhinitis

	Placebo (N=161)	FP200QD (N=158)	LOR10QD (N=163)
Duration of seasonal allergic rhinitis, n(%)			
Less than 2 years	0	0	0
2 to 5 years	20 (12%)	29 (18%)	30 (18%)
6 to 10 years	39 (24%)	33 (21%)	35 (21%)
11 to 15 years	32 (20%)	33 (21%)	31 (19%)
16 to 20 years	13 (8%)	14 (9%)	16 (10%)
21 to 25 years	16 (10%)	12 (8%)	13 (8%)
26 or more years	41 (25%)	37 (23%)	38 (23%)
Duration of seasonal allergic rhinitis (years)			
n	161	158	163
Mean(sd)	17.9(12.7)	16.8(12.4)	16.6(12.6)
Median	13.0	13.0	12.0
Min-Max	2-56	2-51	2-62

Source: Tables 6 from applicant's clinical study report

Disposition of all randomized and treated subjects is shown in Table 9.

Table 9. Disposition of subjects in Study 30034

Demography and Subject Characteristics

	Placebo (N=161)	FP200QD (N=158)	LOR10QD (N=163)	Total (N=482)
Age Group, n(%)				
Less than 18	33 (20%)	29 (18%)	22 (13%)	84 (17%)
18-64	122 (76%)	127 (80%)	134 (82%)	383 (79%)
65-74	5 (3%)	2 (1%)	5 (3%)	12 (2%)
75 or more	1 (<1%)	0	2 (1%)	3 (<1%)
Age (years)				
n	161	158	163	482
Mean(sd)	34.1(15.2)	35.4(14.1)	35.1(14.7)	34.9(14.7)
Median	33.0	37.0	34.0	35.0
Min-Max	12-79	12-69	12-75	12-79
Sex, n(%)				
Female	99 (61%)	106 (67%)	105 (64%)	310 (64%)
Male	62 (39%)	52 (33%)	58 (36%)	172 (36%)
Ethnic Origin, n(%)				
White	103 (64%)	103 (65%)	113 (69%)	319 (66%)
Black	29 (18%)	26 (16%)	26 (16%)	81 (17%)
Asian	2 (1%)	3 (2%)	4 (2%)	9 (2%)
American Hispanic	27 (17%)	25 (16%)	19 (12%)	71 (15%)
Other	0	1 (<1%)	1 (<1%)	2 (<1%)
Height (cm)				
n	161	157	163	481
Mean(sd)	168.1(10.3)	167.3(8.8)	167.4(8.8)	167.6(9.3)
Median	168.0	168.0	166.0	168.0
Min-Max	137-191	145-191	144-193	137-193
Weight (kg)				
n	161	156	163	480
Mean(sd)	78.7(21.0)	75.2(18.7)	77.0(16.9)	77.0(19.0)
Median	75.0	72.5	77.2	75.5
Min-Max	35-168	33-141	42-126	33-168

Source: Table 5 from applicant's clinical study report

The over 90% completion rate, together with the large sample sizes in each group and the averaging over time, obviate the need for detailed considerations regarding missing data.

Statistical Methodologies

The statistical methodologies utilized in Study 30034 are the same as the utilized in Study 30033. See section 3.2.1 for details.

Results and Conclusions

The results for the comparisons of TOSS at each week and overall are shown in Table 10.

Table 10. Results from analyses of TOSS in Study 30034

Change from Baseline in Subject-Rated Total Ocular Symptom Scores

	Placebo (N=161)	FP200QD (N=158)	LOR10QD (N=163)	FP200QD vs Placebo[1]	LOR10QD vs Placebo[1]	FP200QD vs LOR10QD[1]
Baseline[2]						
n	161	158	163			
Mean adj(se)	208.4(4.0)	212.2(4.0)	208.1(3.9)			
Days 1-7						
n	161	158	163	0.026	0.049	0.779
Mean adj(se)	-45.1(5.3)	-61.0(5.4)	-59.0(5.3)			
Days 8-14						
n	161	154	161	0.033	0.211	0.367
Mean adj(se)	-67.5(6.2)	-85.3(6.3)	-77.8(6.2)			
Days 15-21						
n	155	149	158	0.037	0.372	0.226
Mean adj(se)	-84.7(6.7)	-103.6(6.8)	-92.7(6.6)			
Days 22-28						
n	152	143	157	0.151	0.393	0.543
Mean adj(se)	-93.8(6.9)	-107.2(7.1)	-101.6(6.8)			
Days 1-28						
n	161	158	163	0.055	0.216	0.485
Mean adj(se)	-72.0(5.7)	-86.7(5.8)	-81.4(5.7)			

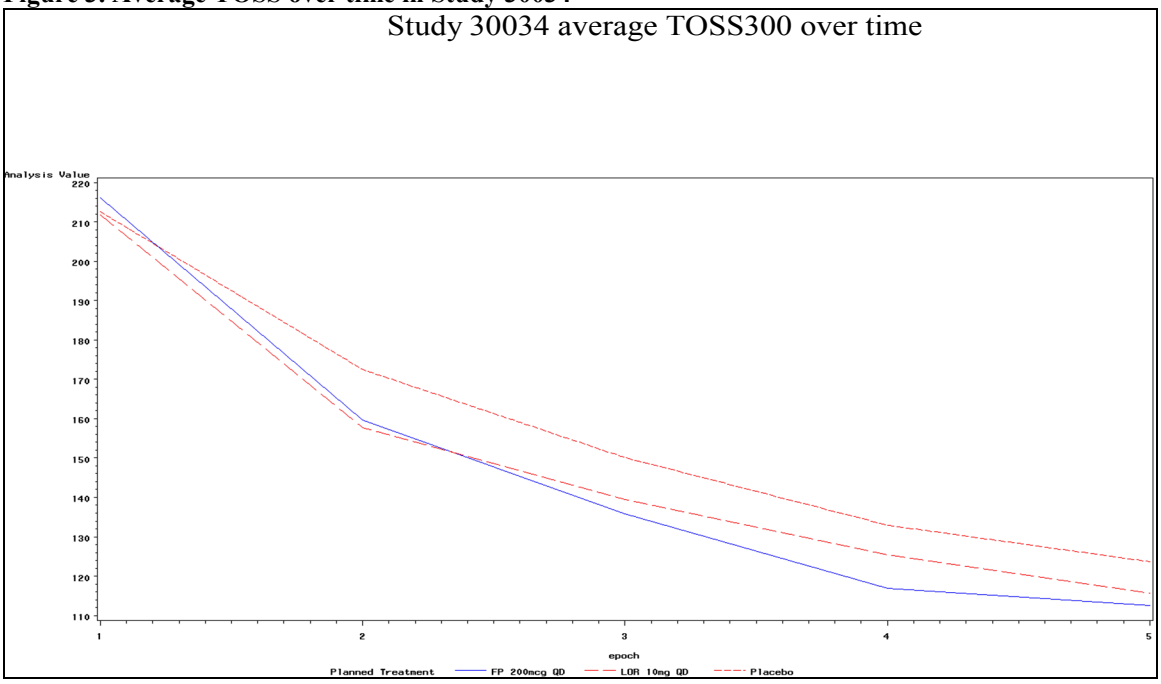
Note: Total Ocular Symptom Score is the sum of itching, tearing and redness scores.
Note: Values shown are actual values at baseline, and change from baseline for subsequent time periods.
[1] p-values are based on analysis of covariance on change from baseline, adjusting for investigator and baseline value, using linear contrasts for pairwise comparisons.
[2] Baseline is defined as the average over the 7 days immediately preceding Day 1 (i.e., treatment start date).

Source: Table 10 from applicant's clinical study report

The marginal p-value of .055 for the primary endpoint is due to the nearly identical mean change from baseline in the FP200QD group to that in Study 30033 ((-88.7 vs -86.7), which is offset by the substantially greater placebo response in Study 30034 (-59.9 vs -72.0). P-values for itching tearing, and redness were .032, .210, and .037, respectively when averaged over days 1-28.

Due to the different placebo responses in the two studies, 30034 was less successful in showing statistically significant treatment differences between placebo and FP200QD with respect to the three symptoms averaged of days 22-28. P-values for itching, tearing, and redness were .10, .29, and .16, respectively. The means of the TOSS over time are shown in Figure 3. Epoch 1=days -6 to 0 (baseline), epoch 2=days 1-7, epoch 3=days 8-14, epoch 4= days 15-21, and epoch 5= days 22-28.

Figure 3. Average TOSS over time in Study 30034

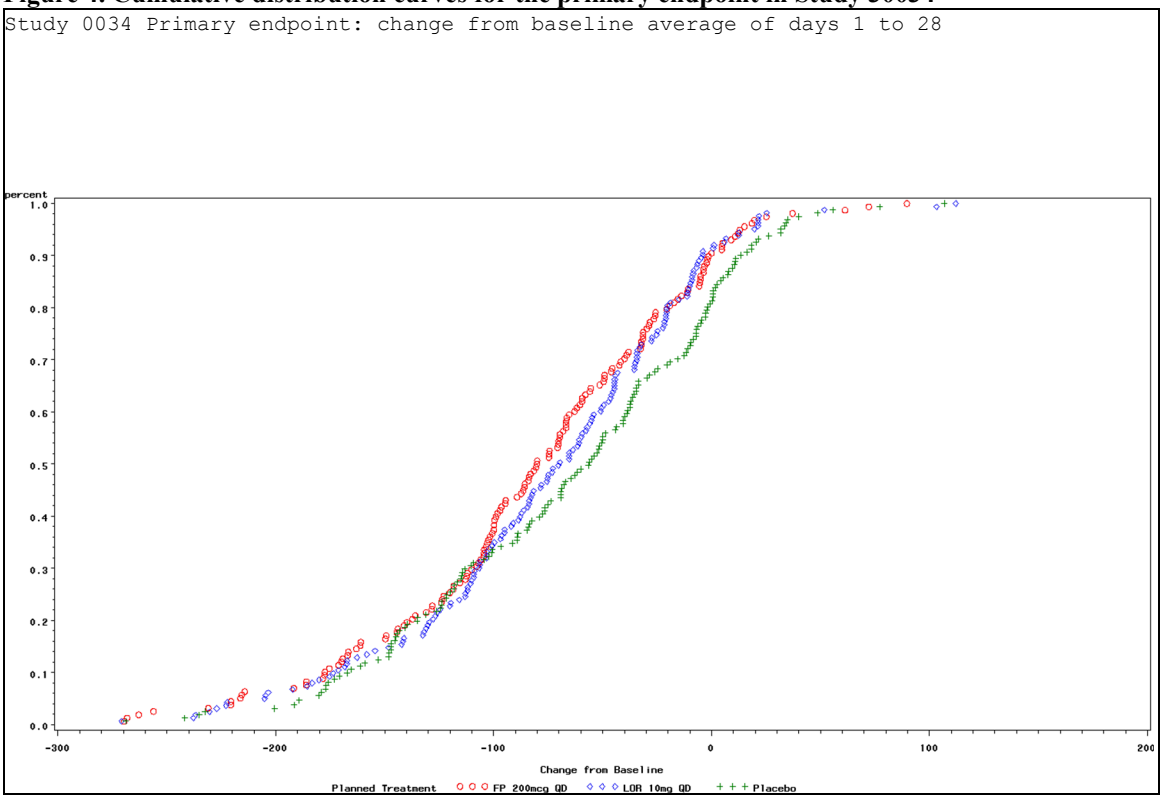


Source: Reviewer

Note the consistent separation between FP200QD and placebo throughout the study duration.

The cumulative distribution curves for the primary endpoint are shown in Figure 4.

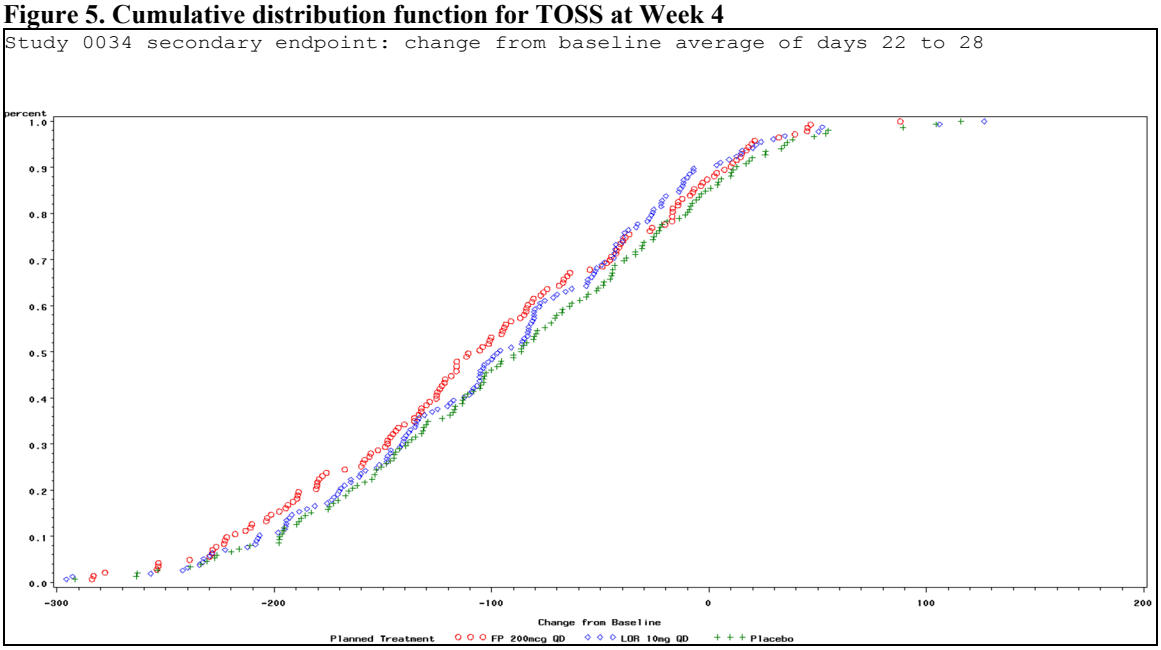
Figure 4. Cumulative distribution curves for the primary endpoint in Study 30034



Source: Reviewer

However, the lack of separation of the average TOSS for days 22-28 can be seen in Figure 5.

Figure 5. Cumulative distribution function for TOSS at Week 4



Source: Reviewer

Trial 30034 failed to confirm trial 30033 with respect to the subject-rated overall evaluation of response to treatment, p-value of .354. Results are shown in Table 11.

Table 11. Results for the analysis of subject-rated overall evaluation of response to treatment in Study 30034

Subject-Rated Overall Evaluation of Response to Therapy

	Placebo (N=161)	FP200QD (N=158)	LOR10QD (N=163)	FP200QD vs Placebo[1]	LOR10QD vs Placebo[1]	FP200QD vs LOR10QD[1]
Evaluation of response, n(%)				0.354	0.772	0.320
n	157	154	161			
Significantly improved	30 (19%)	39 (25%)	34 (21%)			
Moderately improved	39 (25%)	48 (31%)	43 (27%)			
Mildly improved	44 (28%)	33 (21%)	47 (29%)			
No change	30 (19%)	20 (13%)	27 (17%)			
Mildly worse	8 (5%)	5 (3%)	5 (3%)			
Moderately worse	4 (3%)	6 (4%)	5 (3%)			
Significantly worse	2 (1%)	3 (2%)	0			

Source: Table 15 from applicant's clinical study report

In general, Study 30034 did not replicate the unambiguous statistical evidence of efficacy shown in Study 30033.

3.2.3 Study RH01619

This study was of a shorter duration, 2 weeks versus 4 weeks, and used subject-rated reflective total ocular symptom scores instead of total ocular symptom scores and was conducted from December 2012 – February 2013.

Study Design and Endpoints

Study RH01619 was two week study of FP 200QD vs placebo in allergic rhinitis. Only ocular symptoms are examined in this review. The average score over time for the subject-rated reflective total ocular symptom scores (rTOSS) was the primary endpoint and consisted of an assessment of redness, itchiness, and redness. Each symptom was retrospectively rated after 12 hours in the AM and then in the PM. Subjects kept diary scores between 0 (absent)-3(severe) for each of three symptoms: redness, itchiness, and redness. Thus the maximum total score could be 9. The AM and PM scores were averaged. In other words, the AM and PM values were averaged to produce a unique baseline value and unique average endpoint value over time for each subject. Secondary endpoints examined were the mean change from baseline in both the AM and PM rTOSS and the mean change from baseline in the individual symptoms that comprise the rTOSS.

Patient Disposition, Demographic and Baseline Characteristics

Demographics for randomized and treated subjects are shown in Table 12.

Table 12. Patient demographics for Study RH01619

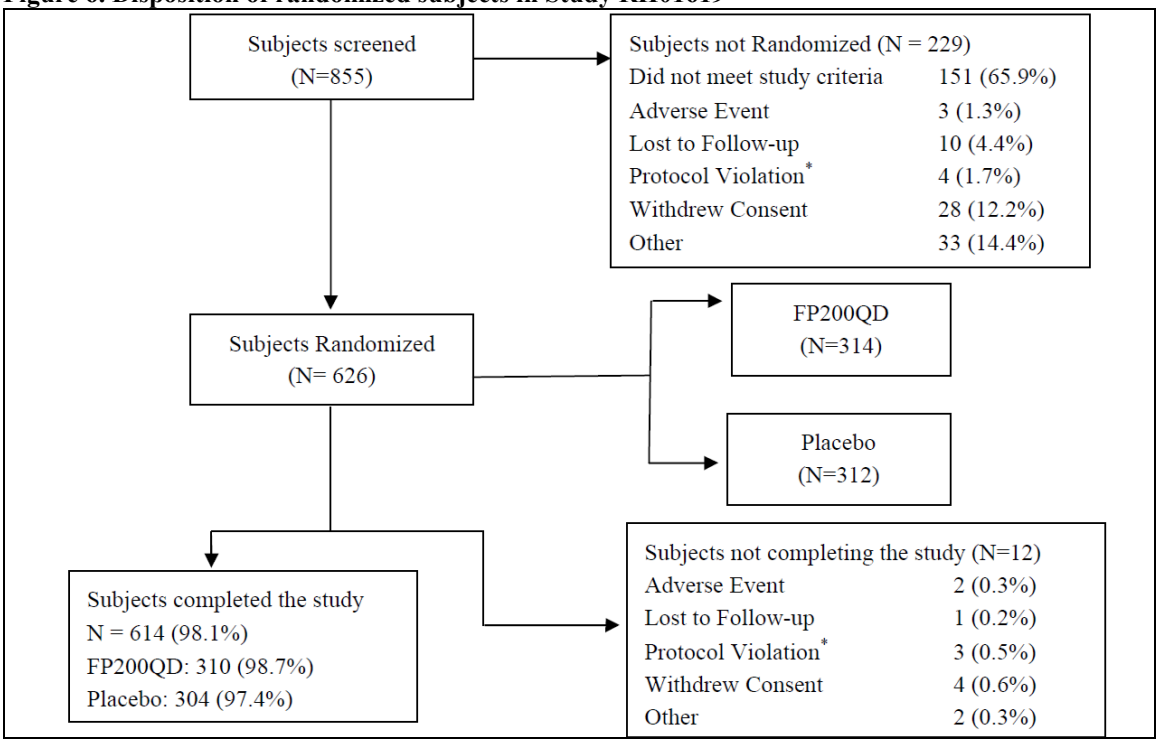
	Placebo (N=312)	FP200QD (N=314)	Overall (N=626)
Sex: N (%)			
Male	111 (35.6)	102 (32.5)	213 (34.0)
Female	201 (64.4)	212 (67.5)	413 (66.0)
Race N(%)			
American Indian or Alaska Native	0	1 (0.3)	1 (0.2)
Asian	2 (0.6)	1 (0.3)	3 (0.5)
Black or African American	38 (12.2)	36 (11.5)	74 (11.8)
White	270 (86.5)	272 (86.6)	542 (86.6)
Native Hawaiian or Other Pacific Islander	1 (0.3)	1 (0.3)	2 (0.3)
Multiple	1 (0.3)	3 (1.0)	4 (0.6)
Race Group: N (%)			
White	270 (86.5)	272 (86.6)	542 (86.6)
Black	38 (12.2)	36 (11.5)	74 (11.8)
Other Race	4 (1.3)	6 (1.9)	10 (1.6)
Ethnicity: N (%)			
Hispanic	124 (39.7)	144 (45.9)	268 (42.8)
Not Hispanic	188 (60.3)	170 (54.1)	358 (57.2)
Age (Years)			
Mean	40.5	40.4	40.5
SD	16.36	14.55	15.47
SE	0.93	0.82	0.62
Median	42.0	41.0	41.0
Minimum	12	12	12
Maximum	79	79	79
Age Group (Years): N (%)			
12-17	47 (15.1)	28 (8.9)	75 (12.0)
18-64	249 (79.8)	273 (86.9)	522 (83.4)
65 or Older	16 (5.1)	13 (4.1)	29 (4.6)

Source: Page 47 from applicant's clinical study report

This study evaluated mainly Caucasian subjects with slightly more female subjects than males but was consistent amongst the three treatment arms. The mean age was approximately 40 years.

Virtually all of the 626 randomized subjects completed the study. Disposition of all randomized subjects is shown in Figure 6.

Figure 6. Disposition of randomized subjects in Study RH01619



Source: Page 46 from applicant's clinical study report

Baseline rTOSS scores for all randomized and treated subjects are shown in Table 13.

Table 13. Baseline rTOSS for subjects randomized in Study RH01619

	Placebo Mean (SD) (N=312)	FP200QD Mean (SD) (N=314)	Overall Mean (SD) (N=626)
TOSS (Maximum Score - 9)			
Daily Reflective	6.98 (1.365)	6.75 (1.364)	6.87 (1.368)
AM Instantaneous	6.89 (1.360)	6.81 (1.399)	6.85 (1.379)
AM Reflective	7.01 (1.391)	6.88 (1.358)	6.95 (1.375)
PM Reflective	6.95 (1.434)	6.63 (1.497)	6.79 (1.474)

Source: CSR, p. 47

As expected, there were no significant differences amongst treatment arms for the baseline rTOSS. Pollen counts by site are shown in Table 14.

Table 14. Mean pollen counts by site in Study RH01619

Site	Placebo run-in Mean (SD)	Treatment Period Mean (SD)
1 (New Braunfels, TX)	578.20 (1261.26)	1590.00 (2520.46)
2 (Austin, TX)	147.60 (366.44)	124.80 (332.21)
3 (San Antonio, TX)	3529.90 (6419.87)	3119.60 (5452.95)
4 (Kerrville, TX)	5297.80 (4351.48)	4784.10 (3628.59)
5 (San Antonio, TX)	4069.00 (7366.01)	3651.00 (6133.38)
6 (Austin, TX)	63.10 (106.11)	159.00 (380.55)

Source: CSR, p. 4

Note that subjects at two sites, Austin and New Braunfels, had minimal exposure to pollen.

Statistical Methodologies

Results for rTOSS were compared between treatment arms using an ANCOVA with treatment, site and baseline rTOSS in the model. Last observation carried forward was used for missing data. However in this study the impact of missing data was minimal since the completion rate was greater 98%. This review reports the average of AM and PM for each of the three symptom scores. There was no attempt to control Type I error for these endpoints.

The sample size was based on the results of a previous study and was designed to show a difference of 0.2 units in mean change from baseline between drug and placebo *for each individual symptom*, not the average of the symptoms which was the primary endpoint. The sponsor states that a difference of 0.6 would be a clinically significant because this difference is found in three studies of a currently marketed fluticasone furate product.

Results and Conclusions

Results of the primary analysis are shown in Table 15.

Table 15. Results from the primary analysis in Study RH01619

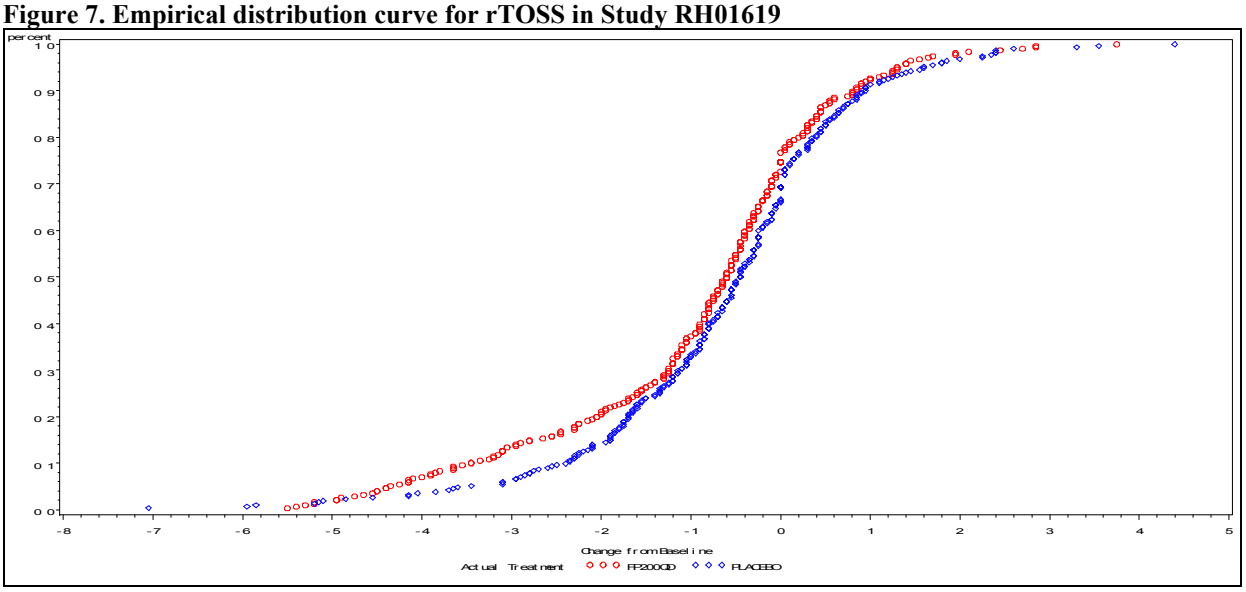
	Placebo (N=312)	FP200QD (N=314)
Baseline		
Mean (SD)	6.98 (1.365)	6.75 (1.364)
Min-Max	3.70 - 9.00	3.60 - 9.00
Post-Baseline (Day 14)		
Mean (SD)	6.35 (1.686)	5.85 (1.767)
Min-Max	0.50 - 9.00	0.30 - 9.00
Change from baseline		
Mean (SD)	-0.63 (1.525)	-0.91 (1.625)
Min-Max	-7.10 - 4.40	-5.50 - 3.80
LS Mean (SE) [1]	-0.61 (0.084)	-0.97 (0.083)
LS Mean Difference [1]		-0.36
95% CI for difference [1]		(-0.59, -0.13)
Between treatment p-value [1]		0.0024

Source: Section 9, Table 9.3.1.1 and Table 9.3.1.2
Baseline score is the average of Days -5 to -1.
Placebo=Placebo nasal spray FP200QD: Fluticasone propionate aqueous nasal spray
[1] From ANCOVA model: Treatment and site as fixed factors and baseline score as covariate. LS mean difference is FP200QD *minus* placebo.

The FP200QD and placebo changes from baseline were- 0.91 and -0.61 , respectively, for a difference of -.36 in favor of FP200QD, well short of the anticipated -0.60 on the 9 point integer scale. This difference produced a p-value of .0024. This reviewer has confirmed this result. One interesting finding is that the low pollen counts at the three of the sites had no modifying effect on the treatment differences. Thus the treatment effect appears to be independent of pollen exposure. Moreover, high or low pollen count did not have an appreciable effect on either site stratum's average change from baseline in average symptom score (-.83 vs -.72).

The cumulative distribution functions of the primary endpoint are shown in Figure 7.

Figure 7. Empirical distribution curve for rTOSS in Study RH01619

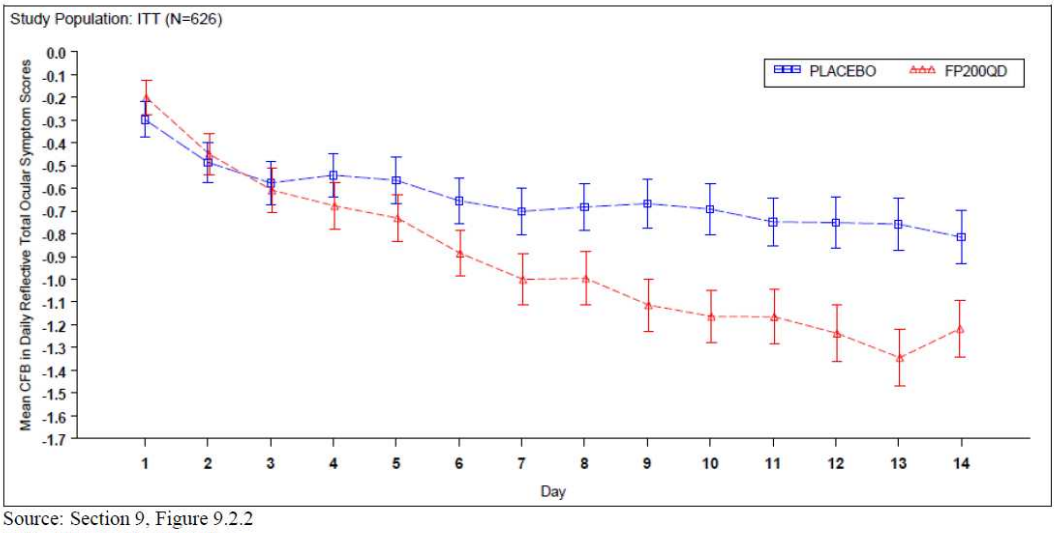


Source: Reviewer

This curve suggests that whatever benefits the drug confers, occurs with *improvement* from baseline but does not produce a benefit among subjects who get worse from baseline. Among those who improved from baseline, there was no clear statistical difference between drug and placebo comparing the sites with the 3 highest (-.36, stderr=.25) and 3 lowest (-.24, stderr=.23) pollen counts during the treatment phase.

Mean rTOSS scores over time are show in Figure 8. Separate plots for AM and PM look similar to the averaged plot shown in Figure 8.

Figure 8. Mean rTOSS scores over time in Study RH01619



Source: Section 9, Figure 9.2.2
CFB - Change from baseline
Placebo=Placebo nasal spray FP200QD: Fluticasone propionate aqueous nasal spray

Source: Page 51 from applicant's clinical study report

Comparisons of the individual symptoms that make up the rTOSS instrument are shown in Table 16.

Table 16. Comparison of individual eye symptoms in Study RH01619

Eye Symptom	Change from Baseline		p-value
	FP200QD	Placebo	
Itching/burning	-.37	-.24	.0023
Tearing/watering	-.35	-.20	.0004
Redness	-.25	-.16	.0319

Source: Reviewer

Note that the differences between the groups are no greater than 0.15 of a unit on a 4-point integer scale. In contrast to the result on redness, the physician's assessment observed no benefit to the drug (p=.86) on a 4-point integer scale.

3.3 Evaluation of Safety

The primary medical officer, Dr. Stacy Chin, reviewed the safety data for this application.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

The primary efficacy endpoint for all studies was examined for a significant treatment interaction with age, racial subgroup. Since these studies were conducted in the United States, no interaction with geographic region was examined. The results for each study are discussed separately below

Study 30033

There was no evidence of treatment by age interaction with age strata younger vs older than 35 years old (p=.80). Nor was there any evidence of treatment by race interaction (p=.82) or gender (p=.32).

Study 30034

There was no evidence of treatment by age interaction with age strata younger vs older than 35 years old (p=.60) and also for treatment by gender (p=.41). Although the interaction p-value for white vs non-white was not significant, the treatment difference for whites was -8 units while that for non-whites was -30 units.

Study RH01619

There appear to be slight more female subjects than male subjects, 66% versus 34%. There was a clear treatment by gender interaction (p=.003) with the lsmean treatment difference within males being .08, and among females, the difference was -0.6. The drug did not confer an average benefit among males.

Using age 41 as the median age for two groups, there is no evidence of treatment by age interaction (p=.96). There are not enough subjects under the age of 18 to assess differential treatment effects between much younger and older cohorts.

Eighty-six percent of the subjects were white. There no evidence of treatment by race interaction (p=.34). However, the treatment difference among non-whites was twice that than among whites (-.69 vs -.31).

4.2 Other Special/Subgroup Populations

There were no other subgroups of interest that were identified or analyzed.

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

There are no significant problems with missing data, especially in Studies 30033 and RH01619, at least in terms of p-values clearly being less than .05. Even Study 30034 had over a 90% completion rate, and it should be noted that the analyzed scores were averages within subjects, thus lessening the influence of any score that was not recorded for any particular day for a given subject. Any imputation technique would not resolve the interpretation of the borderline result of this trial.

5.2 Conclusions and Recommendations

Trial 30033 demonstrated a statistically significant difference between FP200QD and placebo with respect to the primary endpoint, the TOSS300. However, the degree of this difference was not observed in Trial 30034. The marginal result in Trial 30034 (p=.055) was due to a considerably greater placebo response than in Trial 30033. However, the average TOSS scores of the FP200QD groups over 24 weeks in both studies were the same. There was no substantial evidence of treatment by gender interaction or treatment by age interaction.

Trial RH01619's results were somewhat anomalous despite the statistically significant difference between FP200QD and placebo in this two week study. My examination of the empirical distribution functions does not show a clean location shift, there was no relationship between average performance of treatment arm and different pollen counts at the six sites, and on average, males derived no benefit from FP 200 QD over Placebo. The clinical relevance of these findings needs to be evaluated by the clinical review team.

5.3 Label Review

Not applicable.

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

/s/

DAVID HOBERMAN
05/28/2014

DAVID M PETULLO
05/30/2014

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 205-434 **Applicant: GlaxoSmith Kline** **Stamp Date: 9/23/13**

Drug Name: Flonase **NDA/BLA Type:NDA**

On **initial** overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	x			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	x			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).			x	Don't know if applicable
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).	x			

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? ___Yes_____

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	**x**			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	**x**			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.			**x**	
Appropriate references for novel statistical methodology (if present) are included.			**x**	
Safety data organized to permit analyses across clinical trials in the NDA/BLA.				**Don't know**
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.				**Subject to review**

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

David Hoberman	**11/20/13**
Reviewing Statistician	Date
Joan Buenconsejo	11/21/2013
Supervisor/Team Leader	Date

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

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

/s/

DAVID HOBERMAN
11/22/2013

JOAN K BUENCONSEJO
11/22/2013
I concur.

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 205,434 | **Applicant: GSK CH** | **Stamp Date: 9/23/2013**

Drug Name: Fluticasone Propionate Nasal Spray (Flonase Allergy Relief) | **NDA/BLA Type: Standard** | **Indication: Management of symptoms of allergic (b) (4) rhinitis**

This filing checklist pertains only to the consumer studies contained in the submission:

1) A self-selection/actual use study (Study R1810198).titled "An Actual Use Study in Support of the Over-the-Counter Switch of Flonase® Allergy™"
2) A pilot label comprehension study (Study RH01305) titled "Pilot Label Comprehension Study for a Nasal Allergy Treatment"
3) A targeted label comprehension study (RH01318) titled "A Targeted Label Comprehension Study for a Nasal Allergy Treatment"
4) A targeted self-selection study (Study RH01442) titled "A Targeted Self-Selection Study for an Intranasal Allergy Relief Product"
5) Two Human Factors studies
    a. Study RH01801 titled "Usability Test on a Nasal Spray"
    b. Study RH01929 titled "Usability Test on a Nasal Spray in a Low Literacy Population"

On **initial** overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	X			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	X			Study reports are available for all of the consumer studies.
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).			X	
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).	X			Datasets and data definition tables are available for all of the consumer studies

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? ____Yes__

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

Content Parameter (possible review concerns for 74-day letter)	**Yes**	**No**	**NA**	**Comment**
Designs utilized are appropriate for the indications requested.	X			Based on a preliminary review
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	X			Based on a preliminary review
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.			X	
Appropriate references for novel statistical methodology (if present) are included.			X	
Safety data organized to permit analyses across clinical trials in the NDA/BLA.			X	
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.			X	

Reviewing Statistician Date

Supervisor/Team Leader Date

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

/s/

SCOTT S KOMO
11/20/2013

KAREN M HIGGINS
11/20/2013