DEPARTMENT OF HEALTH AND HUMAN SERVICES Public Health Service

Food and Drug Administration
Center for Biologics Evaluation and Research
Office of Biostatistics and Epidemiology
Division of Biostatistics

STATISTICAL REVIEW AND EVALUATION

BLA

BLA/Supplement Number:	125063 / 822
Product Name:	ADVATE (Recombinant Antihemophilic Factor, Plasma / Albumin Free Method)
Indication(s):	Routine prophylaxis in hemophilia A
Applicant:	Baxter
Date submitted:	December 27, 2010
Review Priority:	Standard
Statistical Branch:	Statistical Evaluation
Primary Statistical Reviewer:	John Scott, Ph.D. August 24, 2011
Concurring Reviewer (1):	Jeongsook (Jessica) Kim, Ph.D. August 24, 2011
Concurring Reviewer (2):	Boguang Zhen, Ph.D. August 24, 2011
Medical Office/Division:	OBRR / DH
Clinical Reviewer(s):	L. Ross Pierce, M.D.
Project Manager:	Beth Walton

Table of Contents

1. EXECUTIVE SUMMARY

This is a Prior Approval Supplement for the sponsor's Antihemophilic Factor (Recombinant), Plasma / Albumin-Free Method (ADVATE rAHF-PFM). The product has been licensed in the U.S. since 2003 and is currently indicated for the control and prevention of bleeding episodes in adults and children (0-16 years) with hemophilia A and for peri-operative management in adults and children (0-16 years) with hemophilia A. This supplement contains a final clinical study report and related documents for a Phase IV study investigating the efficacy, pharmacokinetics and safety of two prophylactic regimens of Advate. The sponsor is requesting approval to update the label for Advate to include a new indication for routine prophylaxis to prevent or reduce the frequency of bleeding episodes in adults and children with hemophilia A. The revised label would also include new prophylactic dosage information and a summary of the Phase IV prophylaxis study.

1.1 Conclusions and Recommendations

The statistical evidence supports the sponsor's proposed inclusion of a new indication for routine prophylaxis under either of the two dosing regimens studied in trial 060201. There were three potential statistical points of concern with the sponsor's draft labeling as originally submitted (specifically, with Section 14.4 of the Package Insert, which summarizes the results of the Phase IV prophylaxis study):

1. The sponsor had not reported the result of the comparison between the prophylaxis arms, even though this was the prespecified primary objective of the study.
2. The sponsor had reported annual bleed rates (ABR) in tabular form for the on-demand period, for each prophylaxis arm, and for both prophylaxis arms combined. The ABR is reported separately for all bleeds, for joint bleeds, for non-joint bleeds, for spontaneous bleeds and for traumatic bleeds. While reporting these rates were acceptable, the sponsor had also provided footnotes indicating for which ABR types there was a statistically significant difference between prophylaxis and on-demand regimens. As there was no prespecified method for controlling for multiple comparisons in secondary endpoints, no statistical significance should have been reported for ABR subtypes. Statistical significance should be reported only for ABR for all bleeds.
3. The sponsor had reported statistically significant differences between on-demand and prophylaxis regimens for two SF-36 subscores. Because there was no prespecified method for controlling for multiple comparisons across secondary objectives, no statistical significance should have been claimed for HRQoL endpoints.

After several rounds of information requests and responses from the sponsor, these labeling issues have all been resolved in Amendment 6 of the supplement, as follows:

1. The label now accurately states that no significant difference was found between the two prophylaxis regimens in ABR, with the additional note that the study was not powered to demonstrate equivalence between the two regimens.
2. All claims of statistical significance for specific types of bleeds have been removed.

3. The claims of statistical significance for two SF-36 subscores have been removed. The sponsor has instead included a table summarizing mean differences between regimens in all eight SF-36 domains, along with confidence intervals for these differences.

1.2 Brief Overview of Clinical Studies

The sponsor's proposed new indication for routine prophylaxis is based on the results of clinical study 060201, which was entitled *A Phase 4 Study Comparing Two Prophylactic Regimens In Subjects With Severe Or Moderately Severe Hemophilia A*. Study 060201 was a randomized, multicenter, open label, parallel arm study comparing two prophylactic regimens of Advate following a 6-month lead-in period of on-demand therapy with Advate.

1.3 Major Statistical Issues and Findings

Study 060201 did not show a statistically significant difference in efficacy between the two prophylaxis regimens studied, and both prophylaxis regimens yielded significantly lower annualized bleed rates than were observed during on-demand therapy.

The sponsor also claimed a statistically significant improvement for the prophylaxis vs. on-demand regimens for the SF-36 body pain and physical component scores. However, these comparisons were not included in a comprehensive approach to multiplicity across all secondary objectives, and the claim of statistical significance should not be included in the label.

2. INTRODUCTION

2.1 Overview

This is a Prior Approval Supplement for the sponsor's Antihemophilic Factor (Recombinant), Plasma / Albumin-Free Method (ADVATE rAHF-PFM). The product has been licensed in the U.S. since 2003 and is currently indicated for the control and prevention of bleeding episodes in adults and children (0-16 years) with hemophilia A and for peri-operative management in adults and children (0-16 years) with hemophilia A. This supplement contains a final clinical study report and related documents for a Phase IV study investigating the efficacy, pharmacokinetics and safety of two prophylactic regimens of Advate. The sponsor is requesting approval to update the label for Advate to include a new indication for routine prophylaxis to prevent or reduce the frequency of bleeding episodes in adults and children with hemophilia A. The revised label would also include new prophylactic dosage information and a summary of the Phase IV prophylaxis study.

2.2 Data Sources

This statistical review is based on the clinical overview, final clinical study report and related datasets provided by the sponsor; these documents are located in the EDR at ---(b)(4)--- ---------- . The sponsor has submitted analysis datasets and case report listings as SAS Transport files. The case report data follows the SDTM format; these datasets were validated by Amy Malla prior to the submission being uploaded to the EDR.

I have verified the results of all efficacy analyses reported in the proposed revised label and PI against the data provided by the sponsor with the submission.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

The sponsor's proposed new indication for routine prophylaxis is based on the results of clinical study 060201, which was entitled *A Phase 4 Study Comparing Two Prophylactic Regimens In Subjects With Severe Or Moderately Severe Hemophilia A*.

Study Design and Endpoints

Study 060201 was a randomized, multicenter, open label, parallel arm study comparing two prophylactic regimens of Advate following a 6-month lead-in period of on-demand therapy with Advate. The study enrolled previously treated patients (i.e. at least 150 prior exposure days) with moderately severe to severe hemophilia A (baseline FVIII levels < 2% of normal) between ages 7 and 65. Exclusion criteria included a history of FVIII inhibitors and previous participation in an Advate study.

The primary objective was to compare the rates of bleeding episodes between standard and PK-driven prophylactic treatment regimens. The secondary objectives were:

- To compare the rates of bleeding episodes between on-demand treatment and the prophylactic regimens
- To compare the weight-adjusted consumption of rAHF-PFM between the 2 prophylactic regimens
- To determine the efficacy of rAHF-PFM for the control of bleeding episodes
- To determine the PK parameters for rAHF-PFM utilizing at least 3 lots of the IP
- To determine differences in HRQoL between the 2 prophylactic regimens and changes between the on-demand treatment and prophylactic regimens
- To determine the immunogenicity of rAHF-PFM
- To determine the safety and toxicity of rAHF-PFM

Investigational plan

During the on-demand portion of the study (the first 6 months ± 2 weeks), subjects were treated with recommended dosages of Advate only for bleeding episodes. Following the on-demand portion, subjects were randomized to receive either standard or pharmacokinetic-driven prophylactic regimens for a 12-month period. Randomization was stratified by the number of target joints (0 vs. 1-2 vs. 3 or more) at the end of the on-demand period, where a target joint was defined as a joint in which at least 4 bleeding episodes occurred in the past six months or in which the patient had had at least 20 lifetime bleeding episodes.

Subjects could infuse the defined dose ±20%. Advate was to be administered by intravenous bolus at a maximum infusion rate of 10 mL/minute. Subjects with bleeding episodes and subjects who had undergone a dental or other minor surgery or emergency surgery were to resume their original prophylaxis regimen the next scheduled day after the last therapeutic infusion for the treatment of the bleeding episode or for control of hemostasis following dental, minor and emergency surgery.

For the standard prophylaxis arm, the dose for each subject within the allowable range (20 to 40 IU/kg) was to be determined by the investigator, taking into account the subject's FVIII half-life, incremental recovery, and clinical status. Infusions were to be given every 48 ±6 hours at a routine time determined by the subject and investigator. The dose may have been altered by the investigator within the allowable range depending on clinical circumstances and in accordance with current practice. Any changes in the dose of the standard prophylaxis regimen were to be recorded.

For the PK-driven prophylaxis arm, the exact dose within the allowable range (20 to 80 IU/kg) for each subject was to be determined by the sponsor, utilizing a validated computer program, on the basis of the results of the pharmacokinetic evaluation for each subject. Infusions were to be given every 72 ±6 hours at a routine time determined by the subject and investigator. The half-life and incremental recovery were to be utilized to customize the dose for each subject targeting trough values of at least 1%, if possible. The formula for determining the weight-adjusted dose for PK-driven prophylaxis was:

$$D = (2^{72/t}) / r$$

where D was dose in IU/kg, 72 was the infusion interval in hours, t was the estimated half-life, and r was the incremental recovery for the subject. If the estimated dose was found to be outside the dose range specified (20 to 80 IU/kg), the minimum or maximum of the dose range was to be specified. Dosing for PK-driven prophylaxis for each subject was to be provided to the investigator prior to randomization. The dose was to be maintained throughout the 12-month treatment period.

Study assessments and data collection

Following screening, subjects had a PK assessment visit, followed by assessment visits every three months during the on-demand and prophylactic periods. At the end of the prophylactic period, subjects returned for a final evaluation visit.

The subject or subject's legally authorized representative was to maintain a diary during the subject's participation in the study. The diary was to include the following information:

1. Infusion record (date, time of the infusion, number of units infused, number of vials utilized, lot number)
2. AEs (All AEs that were not intercurrent or preexisting illnesses were to be recorded)
3. Concomitant medications taken (including immunizations)
4. Assessment of hemostatic efficacy for the treatment of bleeding episodes (see Section 9.5.1.5), if applicable
5. Drug accountability (including number of unused vials of IP remaining in the subject's refrigerator)
6. Healthcare resource utilization (for treatments or visits to healthcare professionals not mandated by the protocol):
    a. Physician office visits
    b. Hemophilia treatment site visits
    c. Emergency department visits (reason and number)
    d. Hospitalizations (and associated length of stay)
        i Reason and date of hospitalization
        ii Overall length of stay
        iii Type of ward attended and length of stay in each ward (eg, intensive care unit (ICU), general, other)
    e. Homecare visits (number of homecare visits by type of caregiver)
    f. Type of laboratory tests (eg, specific blood tests, chest X-rays, ECGs, etc.)

Subjects were also to record details regarding all bleeding episodes in their diaries, including severity, cause and anatomical site(s) affected.

Endpoints

The primary efficacy endpoint was the comparison of the yearly transformed rates of bleeding episodes estimated from each of the 1 year prophylaxis regimens. The number of bleeding episodes during the prophylaxis period was to be square root transformed as $X' = \sqrt{X + 0.5}$ and expressed as an annual rate.

The secondary efficacy endpoints were:

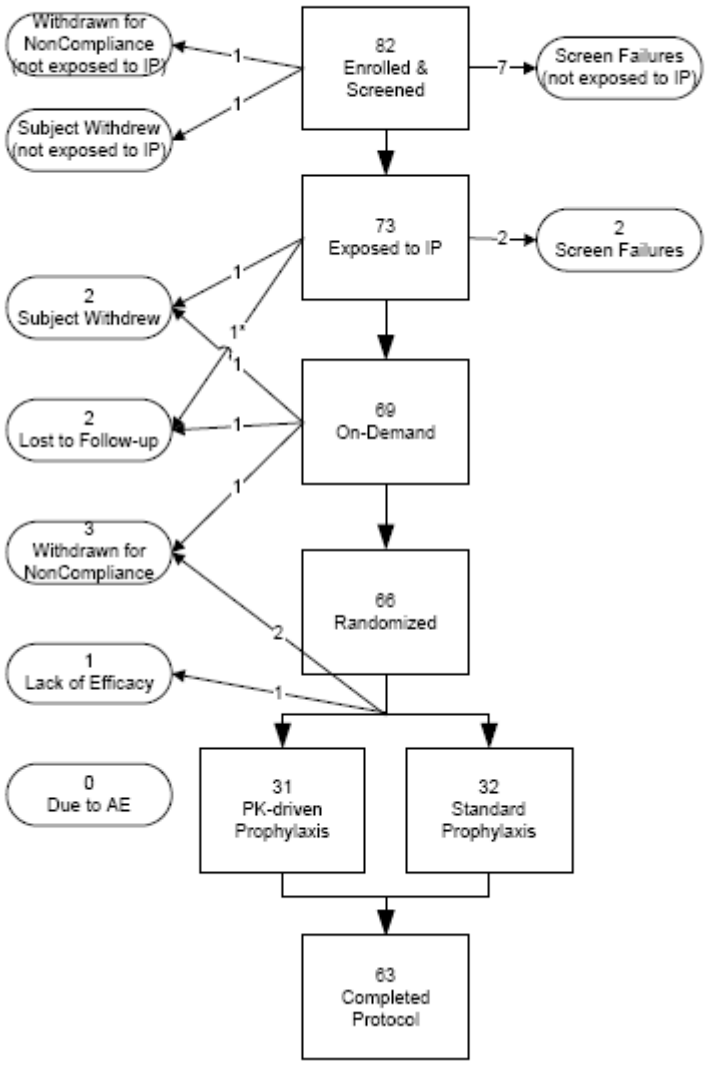
- To compare the yearly transformed rates of bleeding episodes estimated from a 6 month on-demand treatment regimen and each of the 1 year prophylaxis regimens
- To compare the total weight-adjusted units of rAHF-PFM used per year for each of the 1 year prophylaxis regimens

- Bleeding episodes requiring 1, 2, 3, ≥ 4 infusions of rAHF-PFM to achieve adequate hemostasis for each treatment regimen
- Bleeding episodes outcome rated as excellent, good, fair or none for each treatment regimen
- PK parameters estimated for each individual subject

Patient Disposition, Demographic and Baseline Characteristics

A total of 82 subjects were enrolled and screened for eligibility at 21 European and 9 US study sites. Sixty-nine subjects completed on-demand treatment and 67 were randomized to prophylaxis, of which 32 completed PK-driven prophylaxis and 31 completed standard prophylaxis. Reasons for discontinuation are shown in Figure 2.5-1 of the sponsor's clinical overview:

Figure 2.5-1
Participant Flow for Baxter Clinical Study 060201



A total of 1294 protocol deviations were reported during the study for all enrolled subjects, of which 194 were considered by the sponsor to be major deviations. The 194 major protocol deviations occurred in a total of 61 subjects. The majority of these were in the categories of protocol schedule (e.g. visit out of window) or investigational product administration (e.g. incorrect dose administration).

There were several pre-defined analysis sets for the study. Their composition was as follows:

- ITT PK analysis set: included all 71 subjects who provided at least 1 evaluable PK assessment.

- PP PK analysis set: comprised the 63 subjects who provided at least 1 evaluable PK assessment and whose 72-hour PK-driven prophylactic doses were not computed using FVIII activity results from Round Lake central laboratory.
- ITT Efficacy Analysis Set: comprised the 66 subjects in the prophylaxis treatment segment who met predefined criteria of having at least 1 assessment visit or of withdrawing after 3 bleeding episodes prior to the first assessment visit.
- PP Efficacy Set: comprised the 53 subjects who completed both the on-demand and prophylactic treatment periods without any major protocol deviations that would impact the assessment of the primary efficacy endpoint.
- Hemostatic Efficacy Rating Analysis Set: comprised the 71 subjects who reported a bleeding episode on-study that was treated with IP.
- Pharmacoeconomic Analysis Set: comprised the 71 subjects who met the defined criteria (see Section 9.7.1.1.4.1). A total of 66 subjects were included in the pharmacoeconomic analysis subset for those ≥ 14 years of age.
- Safety Analysis Set: comprised all 73 subjects who were exposed to IP.

For randomized subjects who completed at least 1 x3 month interval of prophylaxis treatment but did not complete the full treatment period, the annualized bleed rate was to be estimated from that 1 interval. For subjects who did not have at least 1 x3 month interval during the prophylaxis treatment period, the worst bleed rate observed among subjects who completed at least 1 x3 month interval was to be used. For subjects with missing bleed rates due to surgery, the mean observed rate was to be used.

For the 73 subjects in the safety analysis set, the median age was 26 years, with 6 subjects in the ≥7 and <14 years range, 3 subjects in the ≥14 and <16 range and 64 subjects ≥16 years old. By race category, 64 (88%) of subjects were white, 4 (6%) were Hispanic, 3 (4%) were black or African American, 1 (1%) was Asian and 1 (1%) was other. All subjects were male.

Statistical Methodologies

The primary analysis of the primary efficacy endpoint was a t-test of the null hypothesis that the means of the square root-transformed annualized bleeding rates were equal between the two prophylaxis regimens. The sponsor also performed a Mann-Whitney test comparing the number of bleeding episodes between the two prophylactic regimen as a supportive analysis. The primary efficacy analysis was based on the PP Efficacy Set.

For the secondary endpoint involving the comparison of mean annualized bleeding rates between the on-demand period and prophylaxis period for each of the prophylaxis arms, there was a concern that seasonal effects in bleed rates could confound the analysis, as the subjects were only observed for 6 months during the on-demand portion of the study. The possibility of a seasonal effect in the bleed rate was to be examined using a comparison of median bleed rate for subjects enrolled into the study at various quarters during the year. The possibility of a period effect in bleed rate was to be examined using

a comparison of bleed rate at the first evaluation time point and the second time point just before prophylaxis treatment was initiated. If a change is observed in median bleed rate, this would establish a secular trend. These possibilities were to be examined during the on-demand part of the study. If either a seasonal or secular trend was identified, the comparison of bleed rate between the on-demand treatment regimen and the prophylactic regimen was to use a repeated effects model analysis of variance which incorporated a mixed effect, fixed effect, or covariate parameter, as appropriate. Appropriateness of the statistical model was to be established through examination of the data and the residuals of the fitted model. If no seasonal or period effect was apparent, the comparison of bleed rate between the on-demand treatment regimen and the prophylactic regimen was to use a paired data t-test and Wilcoxon signed rank. The consistency of the enrollment rate was to be examined and presented as a cumulative frequency distribution over the time of enrollment.

The FDA review team felt that a more appropriate approach to assess the possibility of a seasonal effect in the prophylaxis vs. on-demand comparison was to perform a within-subject comparison of the annualized bleed rate between the on-demand period and the corresponding 6 months of the prophylaxis period as a supportive analysis. FDA submitted this request in an IR letter sent to the sponsor in February, 2011, and the sponsor replied with the requested analyses as part of amendment 002 to the supplement, on March 8, 2011.

The comparison of median annual total consumption of rAHF-PFM between the 2 prophylaxis regimens was to utilize a Mann-Whitney test. The median weight-adjusted dose administered per month for the on-demand treatment regimen and each prophylaxis regimen was to be presented.

The frequency distribution for the number of rAHF-PFM infusions required for hemostasis in the treatment of bleeding episodes was to be reported. The mean weight adjusted dose per treatment and per infusion for each bleeding episode were to be calculated.

The sponsor also indicated that Health-related Quality of Life (HRQoL) and healthcare resources utilization data were to be assessed for each treatment regimen. The sponsor indicated that three specific measurement domains of health and an overall health utility index were to be examined, and that all measurements provided in the SF-36 health survey were to be examined and a comparison was to be made between median scores after on-demand treatment and those after prophylaxis treatment using the Wilcoxon signed rank test.

The sponsor's proposed label revision includes a reference to significant differences between on-demand and prophylaxis in SF-36 scores. Because no prospective strategy was provided by the sponsor for multiplicity adjustments across all secondary analyses, and because the HRQoL analyses were not included as secondary efficacy endpoints, these claims should not be included in the label. See Section 4.2 of this review.

The study did not include any planned interim analyses. At the recommendation of the study DSMB, the sponsor performed an unplanned interim analysis of efficacy in late 2008 after data was available for 14 subjects. The sponsor submitted this data as Supplement 562, with the goal of obtaining a revised indication for routine prophylaxis on the basis of the unplanned interim analysis. That supplement was filed with deficiencies, due to Type I error inflation issues related to the unplanned interim analysis for efficacy (see my filing memo for the supplement, dated 2/18/2009). After the mid-cycle review, the sponsor was advised to remove the proposed inclusion of routine prophylaxis from the revised label, and the sponsor did so (see my mid-cycle review memo for the supplement, dated 7/6/2009). Because the interim analysis was not used as the basis for an approval decision, the sponsor has not made any adjustment to the final analyses. This is appropriate.

Results and Conclusions

Primary efficacy analysis

The sponsor's Table 14.2-1 summarizes the annualized bleeding rate data by treatment regimen for the Efficacy PP and ITT sets, and Table 14.2-5 summarizes the results of the primary efficacy comparison between the two prophylaxis arms. The mean square root-transformed annualized bleed rates for the standard and PK-driven prophylaxis arms were 1.46 and 1.61, respectively, for the PP set and 1.58 and 1.90 for the ITT set. There was no statistically significant difference between the arms.

Table 14.2-1
Summary of Annualized Bleeding Rate by Treatment Regimen
(Study 060201: Efficacy Analysis Sets)

Treatment Regimens	Number of Subjects in Treatment Regimen	Number of Subjects With BEs	Number of Bleeding Episodes	Treatment Period in Days Mean (Min, Max)	Annualized Bleeding Rate Mean (SD)	Annualized Bleeding Rate Median (IQR)	Annualized Bleeding Rate (Min, Max)
Per-Protocol							
On-Demand	53	53	1351	185.02 (137, 254)	50.41 (22.23)	43.98 (20.80)	(22.70, 120.51)
Standard Prophylaxis	30	17	77	362.23 (283, 397)	2.56 (4.19)	0.99 (2.14)	(0.00, 17.35)
PK-Driven Prophylaxis	23	14	75	361.48 (286, 382)	3.26 (5.02)	1.00 (4.08)	(0.00, 17.06)
Any Prophylaxis	53	31	152	361.91 (283, 397)	2.86 (4.53)	1.00 (4.07)	(0.00, 17.35)
Intent-to-Treat							
On-Demand	66	66	1640	185.47 (137, 254)	48.95 (21.38)	43.87 (21.92)	(12.97, 120.51)
Standard Prophylaxis	32	19	104	361.59 (283, 397)	3.27 (5.78)	1.00 (3.52)	(0.00, 25.87)
PK-Driven Prophylaxis	34	25	141	344.44 (97, 394)	4.29 (5.13)	2.01 (6.89)	(0.00, 17.06)
Any Prophylaxis	66	44	245	352.76 (97, 397)	3.79 (5.44)	1.10 (4.90)	(0.00, 25.87)

Table 14.2-5
Analysis Results of Primary Efficacy Endpoint: Mean Differences of Transformed Annualized Bleeding Rate Between Each Arm of One Year of Prophylaxis Treatment Regimens
(Study 060201: Efficacy Analysis Sets)

Analysis Set	Transformed Annualized Bleeding Rate PK-Driven Prophylaxis N, Mean	Transformed Annualized Bleeding Rate Standard Prophylaxis N, Mean	Transformed Annualized Bleeding Rate Mean Difference	p-Value[a]
Per-Protocol	23, 1.61	30, 1.46	0.15	0.6016
Intent-to-Treat	34, 1.90	32, 1.58	0.32	0.2588

[a] T-test was used to compare the transformed annualized bleeding episode rates.

Secondary efficacy analyses

For the comparison between on-demand and each prophylaxis arm, the sponsor first evaluated the presence of a seasonal effect by comparing bleed rates during the first three months to the last three months of on-demand treatment, as described above. For the PP efficacy analysis set, the median annualized bleed rate (IQR) during the first half of the on-demand regimen was 47.64 (32.46), compared to 48.16 (29.28) in the second half. Similar results were observed for the ITT efficacy analysis set. The median differences in annual bleed rate between the first half and second half of the on-demand regimen were 0.00 and 1.91 for the PP and ITT efficacy analysis sets, respectively, neither of which were statistically significant. It should be noted that it is not clear that this evaluation

was appropriately designed or adequately powered to detect a seasonal effect in bleed rates.

The sponsor's Table 14.2-8 summarizes the comparison of square root-transformed annualized bleed rates between the on-demand period and each prophylaxis regimen. For the PP efficacy analysis set, the mean difference in transformed annualized bleed rates between subjects in the on-demand regimen and those on standard prophylaxis was 5.42, which was statistically significant ($p < 0.0001$). The mean difference in transformed annualized bleed rates between subjects in the on-demand regimen and those on PKdriven prophylaxis was 5.52, which was also statistically significant ($p < 0.0001$). The mean difference in transformed annualized bleed rates between subjects in the on-demand treatment regimen and those on any prophylaxis regimen was 5.46, which was also statistically significant ($p < 0.0001$). Similar results were obtained using the ITT efficacy analysis set. The magnitude of these effects is such that the differences are extremely unlikely to be attributable to a seasonal effect, even though the sponsor's analysis described in the previous paragraph may not have adequately ruled out the existence of such a seasonal effect.

Table 14.2-8
Analysis Results of Secondary Endpoint: Mean Differences of Transformed Annualized Bleeding Rate Between On-Demand and Prophylaxis Treatment Regimens (Study 060201: Efficacy Analysis Sets)

Treatment Regimen	Number of Subjects	Mean Difference	p-Value[a]
Per-Protocol			
On-Demand and Standard Prophylaxis	30	5.42	<0.0001
On-Demand and PK-Driven Prophylaxis	23	5.52	<0.0001
On-Demand and Any Prophylaxis	53	5.46	<0.0001
Intent-to-Treat			
On-Demand and Standard Prophylaxis	32	5.29	<0.0001
On-Demand and PK-Driven Prophylaxis	34	5.00	<0.0001
On-Demand and Any Prophylaxis	66	5.14	<0.0001

[a] Paired t-test was used to compare the transformed annualized bleeding episode rates.

In Amendment 002, the sponsor provided as supportive analyses within-subject comparisons of the annualized bleed rate between the six months of on-demand treatment and the corresponding six months of prophylaxis treatment. The results of these analyses are provided in Table 15e. The results are qualitatively quite similar to the comparison between on-demand and the full 12 months prophylaxis period reported in Table 14.2-8, supporting the sponsor's conclusion that these results are independent of any seasonal effect.

Table 15e
Mean Differences of Transformed Annualized Bleeding Rate Between On-Demand and the Corresponding 6-month Half of the 12-month Routine Prophylaxis Period (Study 060201: Efficacy Analysis Sets)

Treatment Regimen	Number of Subjects	Mean Difference	p-value[a]
Per-Protocol			
On-Demand and Standard Prophylaxis	30	5.48	<0.0001
On-Demand and PK-Driven Prophylaxis	23	5.59	<0.0001
On-Demand and Any Prophylaxis	53	5.53	<0.0001
Intent-to-Treat			
On-Demand and Standard Prophylaxis	32	5.31	<0.0001
On-Demand and PK-Driven Prophylaxis	31	5.11	<0.0001
On-Demand and Any Prophylaxis	63	5.21	<0.0001

[a] Paired t-test was used to compare the transformed annualized bleeding episode rates.

A comparison of the median annual total consumption (weight-adjusted dose) of Advate between the 2 prophylactic treatment regimens is provided for the PP and ITT efficacy analysis sets in Table 14.2-10. For the PP efficacy analysis set, the median (IQR) annual total weight adjusted dose of Advate was 5768.2 (1778.7) IU/kg for subjects on standard prophylaxis compared to 6118.7 (5795.3) IU/kg for subjects on PK-driven prophylaxis. For the ITT efficacy analysis set, the median (IQR) annual total weight adjusted dose of Advate was 5768.2 (1697.4) IU/kg for subjects on standard prophylaxis compared to 5197.8 (5005.1) IU/kg for subjects on PK-driven prophylaxis. There were no statistically significant differences in median annual total consumption.

Table 14.2-10
Comparison of Median Annual Total Consumption (Weight-Adjusted Dose (IU/kg)) of ADVATE rAHF-PFM between the two Prophylactic Treatment Regimens (Study 060201: Efficacy Analysis Sets)

Analysis Set	PK-Driven Prophylaxis				Standard Prophylaxis				p-Value[a]
	N	Total Weight-Adj Dose (IU/kg)[b]	Median Weight-Adj Dose (IU/kg)[b]	IQR Weight-Adj Dose (IU/kg)[b]	N	Total Weight-Adj Dose (IU/kg)[b]	Median Weight-Adj Dose (IU/kg)[b]	IQR Weight-Adj Dose (IU/kg)[b]	
Per-Protocol	23	139671.5	6118.7	5795.3	30	172427.2	5768.2	1778.7	0.9642
Intent-to-Treat	34	194702.1	5197.8	5005.1	32	183255.7	5768.2	1697.4	0.4924

[a] Mann-Whitney U test.
[b] Annual consumption.

The number of infusions of IP used in the treatment of bleeding episodes is presented by treatment regimen in Table 14.2-11. For the PP efficacy analysis set for all regimens: 1045 (70.8%) bleeding episodes in a total of 49 subjects were treated with 1 infusion; 270 (18.3%) bleeding episodes in 42 subjects were treated with 2 infusions; 111 (7.5%) bleeding episodes in 21 subjects were treated with 3 infusions; and 49 (3.3%) bleeding episodes in 16 subjects were treated with ≥4 infusions. Similar results were observed for the ITT efficacy analysis set.

Table 14.2-11
Number of Infusions Required for the Treatment of Bleeding Episodes by Treatment Regimen
(Study 060201: Efficacy Analysis Sets)

Treatment Regimens	Number of Infusions Required														
	1			2			3			>=4			Total		
	Subj (n)	BE (n)	BE (%)	Subj (n)	BE (n)	BE (%)	Subj (n)	BE (n)	BE (%)	Subj (n)	BE (n)	BE (%)	Subj (n)	BE (n)	BE (%)
Per-Protocol															
On-Demand	49	949	71.1	41	243	18.2	20	105	7.9	15	38	2.8	53	1335	100.0
Standard Prophylaxis	11	49	74.2	5	5	7.6	2	4	6.1	4	8	12.1	16	66	100.0
PK-Driven Prophylaxis	12	47	63.5	5	22	29.7	2	2	2.7	3	3	4.1	14	74	100.0
Any Prophylaxis	23	96	68.6	10	27	19.3	4	6	4.3	7	11	7.9	30	140	100.0
All Treatments	49	1045	70.8	42	270	18.3	21	111	7.5	16	49	3.3	53	1475	100.0
Intent-to-Treat															
On-Demand	62	1168	72.0	51	277	17.1	27	128	7.9	21	50	3.1	66	1623	100.0
Standard Prophylaxis	13	68	73.1	6	12	12.9	2	4	4.3	5	9	9.7	18	93	100.0
PK-Driven Prophylaxis	22	90	64.7	9	37	26.6	4	5	3.6	5	7	5.0	25	139	100.0
Any Prophylaxis	35	158	68.1	15	49	21.1	6	9	3.9	10	16	6.9	43	232	100.0
All Treatments	62	1326	71.5	52	326	17.6	28	137	7.4	23	66	3.6	66	1855	100.0

Abbreviations: Subj (n) = number of unique subjects contributing to the corresponding bleeding episode | BE = Bleeding Episodes.

HRQoL

A summary of the differences in SF-36v1 health domain scores, physical component scores (PCS), and mental component scores (MCS) at the end of each prophylactic treatment regimen compared to the end of the on-demand regimen is presented for the pharmacoeconomic analysis set for subjects ≥ 14 years of age in Table 14.2-18. Due to multiple hypothesis testing across each SF-36v1 scale, adjusted alpha (α*) was set to 0.005 (0.05/10), so a difference was considered statistically significant by the sponsor with a p-value < 0.005. Overall, the differences in health domain scores tended to favor prophylaxis: a negative value for the mean or median difference equates to a larger domain score for the prophylaxis regimen (difference = on-demand – prophylaxis regimen). For 2 domains these differences were statistically significant, both in favor of prophylaxis: the difference in the bodily pain domain between the end of on-demand and the end of any prophylaxis regimen (p =0.0007); and the difference in the PCS between the end of on-demand and the end of any prophylaxis (p =0.0002). The minimal important difference (MID) for bodily pain domain scores over 40 is considered by the sponsor to be 3 points. Therefore the difference in mean bodily pain domain scores between the end of on-demand and any prophylaxis regimen (-4.13) was also considered clinically significant. The MID for mean PCS is considered to be 2 to 3 points. Therefore the difference in mean PCS between the end of on-demand and any prophylaxis regimen (-3.56) was also considered to be clinically significant.

It is important to note that the multiplicity adjustment for the HRQoL analyses do not follow a prespecified strategy that encompasses all secondary efficacy endpoints. Also, HRQoL was not included as a secondary efficacy endpoint in the protocol for this trial, but rather as a secondary objective (the fifth of seven objectives listed). There is

therefore no basis for the sponsor's claim of statistically significant differences in body pain and PCS scores. See Section 4.2 below.

3.2 Evaluation of Safety

Of the 73 subjects in the safety analysis set, 44 (60.27%) reported a total of 200 AEs; 14 SAEs in a total of 11 subjects and 186 non-serious AEs in a total of 41 subjects. There were no deaths or withdrawals due to AEs. A total of 29 subjects in the safety analysis reported no AEs during their participation in this study. Of the 200 AEs, 20 (10.0%) in 4 (5.48%) subjects were judged by the investigator as related to the administration of IP and 180 (90.0%) in 42 (57.53%) subjects were considered as not related. Of the 14 SAEs, 1 (0.5) was considered to be related to the administration of IP. This related mild SAE was a case of a possible low-titer FVIII inhibitor (0.4 BU/mL), which was unconfirmed. Of the 13 unrelated SAEs, 1 (0.5%) was severe in severity, 5 (2.5%) were moderate, 5 (2.5%) were mild, and 1 was not applicable. Of the 19 related non-serious AEs, 3 (1.5%) were moderate in severity and 16 (8.0%) were mild. Of the 167 non-serious unrelated AEs, 4 (2.0%) were severe, 46 (23.0%) were moderate, and 117 (58.5%) were mild.

Further discussion of the safety evaluation is deferred to the clinical reviewer.

3.3 Gender, Race, Age and Other Special/Subgroup Populations

The study enrolled only males, had few pediatric subjects and few non-caucasian subjects. Accordingly, no conclusions can be drawn regarding subgroup-specific differences in treatment efficacy or safety. In Amendment 002, the sponsor provided efficacy and safety results stratified by age and race/ethnicity categories. These tables are not reprinted here**.**

4. SUMMARY AND CONCLUSIONS

4.1 Statistical Issues and Collective Evidence

Study 060201 did not show a statistically significant difference in efficacy between the two prophylaxis regimens studied. However, both prophylaxis regimens yielded significantly lower annualized bleed rates than were observed during on-demand therapy and this evidence could be used to support the proposed new indication for routine prophylaxis under either of the two dosing regimens.

The sponsor also claimed a statistically significant improvement for the prophylaxis vs. on-demand regimens for the SF-36 body pain and physical component scores. However, these comparisons were not included in a comprehensive approach to multiplicity across all secondary objectives, and the claim of statistical significance should not be included in the label.

4.2 Conclusions and Recommendations

The statistical evidence supports the sponsor's proposed inclusion of a new indication for routine prophylaxis under either of the two dosing regimens studied in trial 060201. There were three potential statistical points of concern with the sponsor's draft labeling as originally submitted (specifically, with Section 14.4 of the Package Insert, which summarizes the results of the Phase IV prophylaxis study):

1. The sponsor had not reported the result of the comparison between the prophylaxis arms, even though this was the prespecified primary objective of the study.
2. The sponsor had reported annual bleed rates (ABR) in tabular form for the on-demand period, for each prophylaxis arm, and for both prophylaxis arms combined. The ABR is reported separately for all bleeds, for joint bleeds, for non-joint bleeds, for spontaneous bleeds and for traumatic bleeds. While reporting these rates were acceptable, the sponsor had also provided footnotes indicating for which ABR types there was a statistically significant difference between prophylaxis and on-demand regimens. As there was no prespecified method for controlling for multiple comparisons in secondary endpoints, no statistical significance should have been reported for ABR subtypes. Statistical significance should be reported only for ABR for all bleeds.
3. The sponsor had reported statistically significant differences between on-demand and prophylaxis regimens for two SF-36 subscores. Because there was no prespecified method for controlling for multiple comparisons across secondary objectives, no statistical significance should have been claimed for HRQoL endpoints.

After several rounds of information requests and responses from the sponsor, these labeling issues have all been resolved in Amendment 6 of the supplement, as follows:

1. The label now accurately states that no significant difference was found between the two prophylaxis regimens in ABR, with the additional note that the study was not powered to demonstrate equivalence between the two regimens.
2. All claims of statistical significance for specific types of bleeds have been removed.
3. The claims of statistical significance for two SF-36 subscores have been removed. The sponsor has instead included a table summarizing mean differences between regimens in all eight SF-36 domains, along with confidence intervals for these differences.

DISTRIBUTION LIST

cc:
Original/HFM-380/Beth Walton
DCC/HFM-99
HFM-215 /Chron

HFM-392/L. Ross Pierce
HFM-392/Nisha Jain
HFM-219/Jessica Kim
HFM-219/Boguang Zhen
HFM-215/Christopher Egelebo
HFM-215/Henry Hsu
HFM-215/Estelle Russek-Cohen