CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

761166Orig1s000

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

BLA #:	761166
Supplement #:	Resubmission
Drug Name:	BESREMI (ropeginterferon-a-2b)
Indication(s):	For the treatment of adult patients with polycythemia vera (PV) without symptomatic splenomegaly
Applicant:	PharmaEssentia U.S.A.
Receipt Date:	May 13, 2021
Issue Action Letter:	November 8, 2021
PDUFA Date:	November 13, 2021
Review Priority:	(b) (4)
Biometrics Division:	Division of Biometrics IX
Statistical Reviewer:	Lola Luo
Concurring Reviewers:	Yeh-Fong Chen, Team Leader Thomas E. Gwise, Division Director
Medical Division:	Division of Non-malignant Hematology
Clinical Team:	Patricia Oneal, Medical Reviewer Tanya Wroblewski, Team Leader Ann Farrell, Division Director
Project Manager:	Carleveva Thompson

Keywords: Link to keywords:
http://intranetapps.fda.gov/scripts/ob_apps/ob/eWork/uploads/eWork/2009/Keywords-in-DFS.htm

Since the Sponsor did not make any modification to the efficacy data in this resubmission, there is no additional statistical review needed. The statistical review for the original submission can be viewed in DARRTs (SD#1 and Check in Date: 02/26/2021).

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

LOLA LUO
10/04/2021 03:37:43 PM

YEH FONG CHEN
10/04/2021 03:44:50 PM

THOMAS E GWISE
10/04/2021 04:06:11 PM

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

BLA #:	761166
Supplement #:	001
Drug Name:	BESREMI (ropeginterferon-a-2b)
Indication(s):	
Applicant:	Pharma Essentia Inc.
Receipt Date:	March 13, 2020
PDUFA Date:	March 13, 2021
Review Priority:	Standard
Biometrics Division:	Division of Biometrics IX
Statistical Reviewer:	Lola Luo
Concurring Reviewers:	Yeh-Fong Chen, Team Leader Thomas E. Gwise, Division Director
Medical Division:	Division of Hematology Product
Clinical Team:	Patricia Oneal, Medical Reviewer Tanya Wroblewski, Team Leader Ann Farrell, Division Director
Project Manager:	Carleveva Thompson
Keywords:	

Link to keywords:
http://intranetapps.fda.gov/scripts/ob_apps/ob/eWork/uploads/eWork/2009/Keywords-in-DFS.htm

Table of Contents

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

On 13 March 2020, the applicant, PharmaEssentia U.S.A. LLC, submitted a new drug application (BLA) 761166 for the study drug, Besremi (AOP2014), for the treatment of adult patients with Polycythemia Vera (PV) without symptomatic splenomegaly in the United States.

This BLA includes two phase III clinical studies (PROUD-PV and CONTINUATION-PV) and one phase I/II clinical study (PEGINVERA). The PROUD-PV (pivotal study), was a randomized, open-label, multicenter, active-controlled, parallel arm study to assess the efficacy and safety of AOP2014 in patients with PV. The study was initiated on 04 October 2013 and completed on 08 April 2016. CONTINUATION-PV was an open-label, extension study of PROUD-PV that consisted of patients who had been previously treated in PROUD-PV. CONTINUATION-PV was intended to assess the long-term efficacy of AOP2014 and was initiated on 25 November 2014 and is ongoing, the database lock for the interim analysis was on 29 May 2018. PEGINVERA was an open-label, dose finding study conducted in two stages.

In PROUD-PV study, patients with PV were randomized 1:1 to either the study drug, AOP2014, or the control drug, Hydroxyurea (HU). The primary objective is to show that AOP2014 was non-inferior to HU, in the primary endpoint, disease response rate at month 12, with a proposed non-inferiority margin of 10.5%. A patient was considered to achieve a response if the following 4 criteria were met:

- Hematocrit < 45% without phlebotomy (>3 months since last phlebotomy)
- Platelets < 400 x 10^9/L
- Leukocytes < 10 x10^9/L
- Normal Spleen Size (< 12 cm for females, < 13 cm for males)

The non-inferiority of AOP2014 to HU in the disease response rate can be claimed if the lower bound of 95% CI of the difference in the disease response rate (AOP2014-HU) is greater than the NI margin of -10.5%.

There were several statistical issues with the design and analysis of PROUD-PV. PROUD-PV was originally designed to demonstrate a superiority of AOP2014 compared to HU. In the amended clinical study protocol version 4.0, dated June 15, 2016, sponsor changed the superiority comparison to a non-inferiority comparison. This change was made 2 months after the completion of PROUD-PV (08 April 2016). Given that the study was open-label and the sponsor had access to the unblinded data during the study, it was possible for the sponsor to make changes based on the failed results observed from the superiority comparison, which could violate the integrity of the study. In addition, the NI margin of 10.5% was not adequately justified. In a later meeting, the sponsor further proposed to (b) (4) At the time of these proposals, the review team did not agree to these changes (meeting minutes dated 20 April 2017, 11 April 2019, 04 September 2019).

Two-hundred and fifty-seven patients were randomized into PROUD-PV with 127 patients in the AOP2014 treatment arm and 130 patients in the HU treatment arm. Three patients in HU arm never received the drug and excluded from the primary analysis. There were 217 out of 254

patients (85.4%) completed the study. The difference in response rate was -6.3% (95% CI: -16.7, 4.1). Since the lower bound of the 95% CI was -16.7% and is less than the NI margin of -10.5%, the non-inferiority of AOP2014 to HU at month 12 was not demonstrated.

A possible reason for PROUD-PV to fail could be the different dose titration between the two treatment arms with a very conservative one for the AOP2014 arm. The dose escalation for AOP2014 was deliberately conservative to avoid toxicities and for tolerability. The difference in dose titration resulted in a 20-week delay in reaching the maximum dose for the AOP2014 arm. Given the duration of PROUD-PV was only 12 months, the assessment of efficacy was conducted too early.

CONTINUATION-PV was designed as an exploratory study to assess the long-term efficacy of AOP2014. The study was originally designed as a single-arm extension study only for patients receiving AOP2014. A control arm, best available therapy (BAT), was added later. Patients who received AOP2014 or HU, completed the pivotal study (PROUD-PV), and met the eligibility criteria were enrolled in this study. Patients who received AOP2014 in PROUD-PV were assigned to the AOP2014 arm and those HU patients who enrolled were allocated to the BAT arm (i.e., all patients in BAT arm were assigned HU). There were no power calculation and no formal hypothesis planned; All statistical analyses were performed for exploratory purposes.

Out of 217 patients completed PROUD-PV, 171 patients (95 patients in AOP2014 arm and 76 patients in BAT arm) entered in CONTINUATION-PV. Data were analyzed and reported after Month 24 and 36. CONTINUATION-PV has two co-primary endpoints, which are two different definitions of the disease response:

1. hematologic response, and normal spleen size
2. hematologic response and clinical improvement in disease burden (resolution or clinically improvement of disease-related signs (clinically significant splenomegaly and disease related-symptoms (microvascular disturbances, pruritus, headache).

For the first co-primary endpoint, the relative risk between AOP2014 and BAT was 1.2 (95% CI: 0.7, 1.7) at Month 24 and 1.4 (95% CI: 0.9, 2.1) at Month 36. For the second co-primary endpoint, the relative risk between AOP2014 and BAT was 1.4 (95% CI: 0.9, 1.9) at Month 24 and 1.4 (95% CI: 1.0, 2.0) at Month 36.

CONTINUATION-PV also had several statistical issues. The patient population was consisted of a subgroup of patients who completed the PROUD-PV study and met certain outcome dependent eligibility criteria. This was not a randomized study. Without proper randomization, selection bias may exist and confounders (either known or unknown) could undermine the study results and increase the type I error rate. Further, the study did not have a formal statistical hypothesis or pre-specified decision criterion planned to declare the success of the study.

The single arm data of AOP2014 (n=127) from PROUD-PV/CONTINUATION study at 24- and 36-month were assessed to further assess the efficacy of AOP2014. The single arm data allow for assessment of responses beyond 12 months for CHR, which was defined as HCT < 45%

without phlebotomy (> 3 months since last phlebotomy), platelets, 400 x 10^9/L, leukocytes < 10 x 10^9/L.

Supportive data from the single-arm experience from the PROUD-PV/CONTINUATION study demonstrated a CHR of 52.8% at month 24 and 36 (95% CI of 44.1, 61.4) with median duration of response beyond 48 months.

The PEGINVERA study was also reviewed and the study results showed evidence to support the effectiveness of AOP2014. PEGINVERA was an open-label, dose finding study conducted in two stages. The primary objective of stage I was to identify the maximum tolerated dose (MTD) and Stage II was to assess efficacy, safety and tolerability. Efficacy criteria was based on hematological response defined as:

- HCT < 45% (without phlebotomy in previous two months)
- Platelet count ≤ 400 x10^9/L,
- WBC ≤10 x 10^9/L
- Normal spleen size
- Absence of thromboembolic events

The complete hematological response was 52.3% (95% CI: 38.6, 66.0). The CHR (laboratory measures only) was 80.4% (95% CI: 66.9, 90.2).

Based on the data submitted, this reviewer concludes that the pivotal study, PROUD-PV, failed to demonstrate non-inferiority of study drug, AOP2014, to the active control, HU, in the primary endpoint, disease response rate. However, this reviewer acknowledges that the possible reason for the study failure could be due to the differences in dose titration schedule for the two treatment arms and the planned assessment of efficacy was too early. The CONTINUATION-PV study was designed as an extension study that offered only exploratory comparison between AOP2014 arm to BAT arm. The results from PEGINVERA and the evaluation of AOP2014 as single-arm from combined PROUD-PV/CONTINUATION study appeared to provide some evidence of the effectiveness for AOP2014 given patients with PV do not have spontaneous remission when not treated. However, there may exist some limitations to the statistical interpretability of the efficacy findings in PEGINVERA due to its single arm design, and the post-hoc analysis nature of the single arm of combined PROUD-PV/CONTINUATION study.

Whether a substantial evidence of effectiveness of AOP2014 has been demonstrated in this application need clinical judgment and should be considering unmet medical need.

INTRODUCTION

1.1 Overview

Investigational drug and intended indication

Besremi (Ropeginterferon alfa-2b) is an innovative pegylated interferon with an orphan designation, (granted April 2, 2012), to treat the rare disease PV. Currently, PV patients have limited approved therapeutic options and there is a critical unmet medical need for patient with PV in early disease stage and before failing the hydroxyurea in the United States. Ropeginterferon alfa-2b would address this need with improved patient compliance and a reduction of treatment-limiting reactions due to a lower frequency of injection schedule than other interferons drug product, historically used off-label. Ropeginterferon alfa-2b drug product has been approved in Europe by the European Commission for the treatment of patients with PV and has shown durable clinical benefit with notable tolerability and safety.

For this submission, the applicant proposed a new indication to use Besremi for the treatment of *"P*olycythemia Vera in adults without symptomatic splenomegaly.*"*

Clinical Study

PROUD-PV (Pivotal Study)

- A randomized, open-label, multicenter, controlled, parallel arm, phase III study assessing the efficacy and safety of AOP2014 vs. Hydroxyurea in patients with Polycythemia Vera
- Study started on 04 October 2013 and ended 08 April 2016.
- The study protocol version 4.0 changed the primary objective of the study from demonstrating superiority of AOP2014 above HU (in terms of disease response rate in both HU naïve and currently treated patients, diagnosed with PV to demonstrating non-inferiority.)

CONTINUATION-PV (Extension Study)

- An open-label, multicenter, phase IIIb study assessing the long-term efficacy and safety of AOP2014 and standard first line treatment (BAT) in patients with polycythemia vera who previously participated in the PROUD-PV Study.
- Study started on 25 November 2014. The study is ongoing. Database lock for interim analysis was on 29 May 2018.

PEGINVERA

- An open-label, prespective, multicenter, phase I/II dose escalations study to determine the maximum tolerated dose and to assess the safety and efficacy of AOP2014 in patients with PV.
- Study started on 13 September 2010 and ended 25January 2018.

1.2 Data Sources

Data were provided electronically with the standard analysis data formats. SAS programs used to create key efficacy and safety outputs for the study were submitted in a CD form with this application.

The link of this submission and the data is: \\CDSESUB1\evsprod\BLA761166\0001

1.3 Overview of Polycythemia Vera and Unmet Medical Need

Polycythemia Vera (PV, Polycythemia Rubra Vera, Maladie de Vaquez) is one of the chronic myeloproliferative neoplasms (MPNs). PV is a rare clonal hematopoietic stem cell disorder that displays terminal hematopoietic cell expansion in the peripheral blood resulting in erythrocytosis, leukocytosis, and thrombocytosis. Polycythemia Vera can occur at all ages and more commonly affects women at a younger age but increases in frequency exponentially over the age of 60 with a male predominance. Overall, the median age of diagnosis is approximately 60 years and approximately 25% of the cases present before age 50 and one tenth before the age of 40 years. The prevalence is approximately 50 per 100,000 in the United States.

Median survival for patients with PV is 13.7 years and the following figure depicts overall survival for the myeloproliferative disorders with the green line depicting survival for patients with PV (Blood 2014;124(16): 2507-2513).

PV is a long-term debilitating and life-threating disorder with no spontaneous remissions. The key clinical features of PV generally include elevated hematocrit, increased leukocytes, elevated platelet counts and in more advanced cases splenomegaly. Mild clinical symptoms of PV can be unspecific and likely related to increased blood cell counts and viscosity and include headache, fatigue, dizziness, vision disturbances, vertigo, tinnitus, pruritus and erythromelalgia. Significant morbidity and mortality of the disease is due to the risk of thrombosis (arterial and venous), and the long-term risk of leukemic transformation.

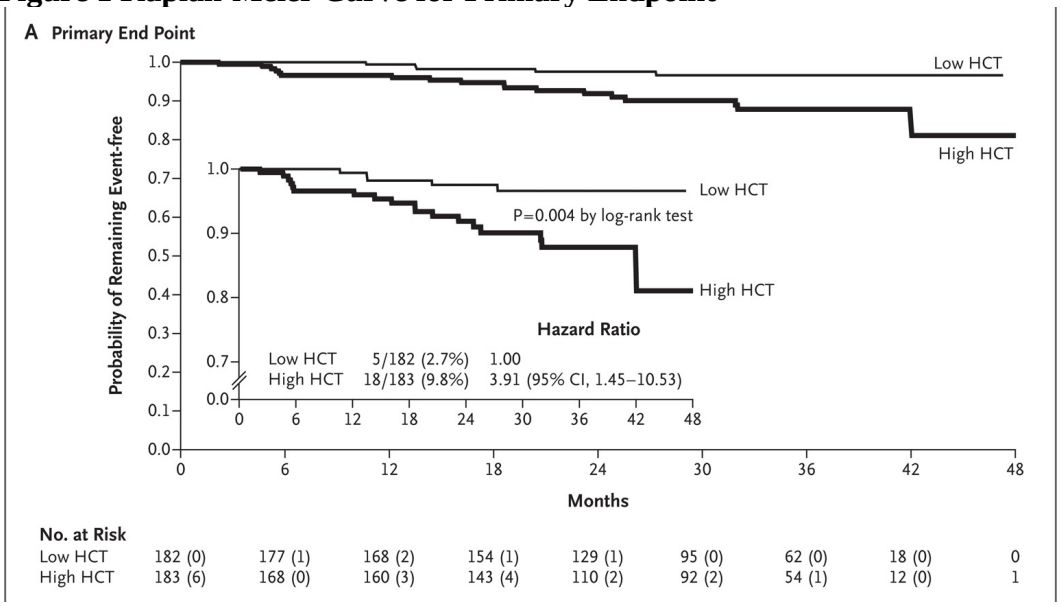
In the past 10-20 years, significant progress has been made in elucidating the underlying molecular pathogenesis of PV. The discovery of the gain of function mutation in the tyrosine kinase (Janus-activated kinase-JAK2V617F) which now appears to be associated with primary PV cases in adults. JAK2V617F is detectable in more than 95% of patients diagnosed with PV. JAK2 mutations cause the enzyme to be constitutively active and leads to an “always on” signal for cell propagation. There is a consensus that the presence of higher allelic burden is associated with clinical markers indicative of a more aggressive phenotype such as higher WBC and splenomegaly.

The therapeutic goal for patients with PV is cytoreduction and to reduce the risk of thrombotic/cardiovascular events. Additional goals include alleviation of symptoms and prolonging survival by decreasing or minimizing the risk of progression to myelofibrosis (MF), or acute leukemia. Current available treatment options are based on risk stratification of the patient. Low risk patients (age < 60, no history of thrombosis) usually receive low dose aspirin and phlebotomy, high risk patients (age ≥60 and prior thrombosis) receive either off-label

hydroxyurea or off-label interferon. Patients who are HU-refractory or intolerant, massive splenomegaly or intractable symptoms usually receive JAK 2 inhibitors such as ruxolitinib, or off-label busulfan, or 32P.

Cytoreduction is a critical aspect in this disease. Elevation of counts increase the risk of thrombosis and death for example, WBC > 11 x 10^9/L is associated with increased risk of death and thrombosis in patients with PV. Targeting hematocrits (HCT) <45% results in lower rate of cardiovascular death and major thrombosis in patients with PV. A prospective study in 365 adults with JAK2 positive PV randomized patients to either a low target HCT (<45%) or high HCT (45-50%). (N Engl J Med 2013; 368:22-33). The primary composite endpoint was time until death from cardiovascular events (CV), CV hospitalizations, incidence of cancer, progression of myelofibrosis, MDS or leukemic transformation, hemorrhage. The following figure is a Kaplan-Meier Curve for death from CV causes or thrombotic events for the primary endpoint demonstrating an improvement in the primary endpoint for the lower HCT group.

Figure 1 Kaplan-Meier Curve for Primary Endpoint



There are no spontaneous remissions in this disease. Phlebotomy and low dose aspirin are the cornerstones of treatment in the initial phase of treatment. Hydroxyurea (HU), which is not FDA-approved for the treatment of PV, is often used as first line cytoreductive therapy in patients with PV; however, HU related toxicities such as fatigue often require drug reduction or discontinuation. Reports of the incidence of leukemic transformation in patients treated with HU have ranged from 1 to 17 percent with lower values obtained at shorter follow-up times. Interferons (pegylated and non-pegylated) have been used for over 30 years in MPN treatment including PV. There are over 40 published studies (single-arm, retrospective in > 1000 patients with PV or PV and Essential Thrombocytosis (ET) describing the effectiveness of interferon (IFN) in the treatment of PV. The published complete hematological remission is reported

between 39-76% in patients with PV with the use of interferons. The most common class adverse reactions of interferon alfa include flu-like symptoms, fatigue, depression (suicidal ideation), diarrhea and weakness.

Ropeginterferon is a type I interferon alpha and. Type I interferon alfa exhibit cellular effects by binding to transmembrane receptor-interferon alfa receptor (IFNAR). Binding to IFNAR initiates downstream signaling cascade through activation of kinases and in particular Janus Kinase 1 and tyrosine kinase 2 and STAT proteins.

PharmaEssentia has developed ropeginterferon alfa-2b (AOP2014) as a novel, more stable third generation pegylated interferon analog containing a mono-pegylated interferon. In contrast to conventional pegylated interferons currently approved in the United States (PEGASYS, PEGINTRON, SYLATRON), which require administration once per week, ropeginterferon alfa 2b can be administered every 2 weeks in patients with PV.

STATISTICAL EVALUATION

1.4 Data and Analysis Quality

Data of this submission, provided with SDTM and ADaM, are acceptable. The sponsor also provided clear definition file for datasets and, reviewer guide and detailed analysis programs for assisting review.

1.5 Evaluation of Efficacy

1.5.1 PROUD-PV (Pivotal Study)

1.5.1.1 Study Design and Endpoints

Study PROUD-PV was a randomized, open-label, multicenter, controlled, parallel arm, phase III study assessing the efficacy and safety of AOP2014 vs. Hydroxyurea (HU) in patients with Polycythemia Vera (PV).

A double-blind design was not an option due to obvious differences in routes of administration (subcutaneous vs. oral) and toxicity between study regimens.

Since the study was conducted open label, the study endpoints could be objective to potential bias: relevant blood values at different timepoints were measured by a central laboratory, blinded to treatment assignment; and an observer-independent imaging (CT/MRI) with blinded radiologic assessment was used for evaluation of spleen size.

Primary Efficacy Objective:

- To demonstrate non-inferiority of AOP2014 vs. HU in terms of disease response rate in both naïve and currently treated patients, diagnosed with PV.

The final analysis of this study included 12-month response data. One formal interim analysis (IA) was planned on 6-month response data.

Selection of Study Population:
Patients who met all the inclusion and none of the exclusion criteria listed in below were invited to participate in the study.

Inclusion Criteria:

1. Male or female, 18 years or older.
2. Diagnosis of PV according to the World Health Organization (WHO) 2008 criteria with the mandatory presence of JAK2V617F mutation as the major disease criterion.
3. For previously cytoreduction untreated patients – documented need for cytoreductive treatment (one or more of the following criteria):

    a. Age >60 years at the planned day of the first drug administration;
    b. At least one previous well documented major cardiovascular PV-related event, except bleeding and PV-related thromboembolic complications in the abdominal area, see excl. criterion 7) in the medical history;
    c. Poor tolerance (defined as a phlebotomy/ procedure-related AE causing significant adverse impact on the patient and limiting ability to apply phlebotomy with the intention to keep Hct <45%) or frequent need for phlebotomy (more than one phlebotomy within last three months prior entering the study, while each of these phlebotomies was performed to reduce Hct level from >45%, or if one phlebotomy was not able to reduce Hct level to <45% for the next three months following phlebotomy);
    d. Progressive splenomegaly (de novo appearance of a palpable spleen, or appearance of the symptoms, related to the enlarged spleen, e.g. pain, early satiety etc., with confirmed size increase);
    e. Platelet counts greater than 1000 x 10^9/L (for two measurements within one week);
    f. Leukocytosis (white blood cell [WBC] >10 x 10^9/L for two measurements within one week).

4. For patients currently treated or pre-treated with HU, all of the following criteria:
    a. being non-responders (as defined by the response criteria for primary endpoint in this protocol);
    b. Total HU treatment duration shorter than three years;
    c. No documented resistance or intolerance as defined by modified Barosi et al, 2009 criteria.

5. HADS score 0-7 on both subscales

6. Patients with a HADS score of 8-10 inclusive on either or both of the subscales may have been eligible following psychiatric assessment that excluded clinical significance of the observed symptoms in the context of potential treatment with an IFN-α.

7. Signed written IC.

Exclusion Criteria:

A patient who met any of the following criteria did not qualify for entry into this trial:

1. Any systemic cytoreduction for PV in the medical history prior to study entry with exception of HU for shorter than 3 years (see respective inclusion criterion).
2. Any contraindication to any of the investigational medicinal products (IMPs) (pegylated IFN or HU) or their excipients.
3. Any systemic exposure to a non-pegylated or pegylated IFN-α in the medical history.
4. Documented autoimmune disease at screening or in the medical history.
5. Clinically relevant pulmonary infiltrates, pneumonia, and pneumonitis at screening.
6. Infections with systemic manifestations, e.g., hepatitis B, hepatitis C, or human immunodeficiency virus (HIV) at screening.
7. Known, PV-related thromboembolic complications in the abdominal area (e.g. portal vein thrombosis, Budd-Chiari syndrome) and/or splenectomy in the medical history.
8. Any investigational drug less than 6 weeks prior to the first dose of study drug or not recovered from effects of prior administration of any investigational agent.
9. History or presence of depression requiring treatment with antidepressant.
10. HADS score equal to or above 11 on either or both of the subscales.
11. Any risk of suicide at screening or previous suicide attempts.
12. Any significant morbidity or abnormality which may interfere with the study participation.
13. Pregnancy and breast-feeding females of reproductive potential and males not using effective means of contraception. Note: women of childbearing potential not using effective contraceptive methods were not eligible for the study. A woman of childbearing potential was defined as any female having experienced menarche and who is not postmenopausal or permanently sterilized (e.g. tubal occlusion, hysterectomy, bilateral salpingectomy) (according to MHRA guidance, see references).
14. History of active substance or alcohol abuse within the last year.
15. Evidence of severe retinopathy (e.g. cytomegalovirus retinitis, macular degeneration) or clinically relevant ophthalmological disorder (due to diabetes mellitus or hypertension).
16. Thyroid dysfunction (clinical symptoms of thyroid hyper- or hypofunction) not adequately controlled.
17. Patients tested positively to thyroglobulin (TgAb) autoantibodies and / or thyroid peroxidase (TPOAb) autoantibodies at screening.
18. History of major organ transplantation.
19. History of uncontrolled severe seizure disorder.
20. Leukocytopenia at the time of screening (leukocytes below the lower limit of normal).
21. Thrombocytopenia at the time of screening (platelets below the lower limit of normal).
22. History of malignant disease, including solid tumors and hematological malignancies (except basal cell and squamous cell carcinomas of the skin and carcinoma *in situ* of the cervix that have been completely excised and are considered cured) within the last 3 years.

Study Treatment:

AOP2014 (Peg-P-IFN-α-2b)

AOP2014 was supplied in 1 ml vials containing 0.5 mg/ml solution. Study drug was given s.c. Inactive ingredients include sodium chloride, polysorbate 80, benzyl alcohol, sodium acetate, and acetic acid. The solution was colorless to light yellow.

AOP2014 was administered s.c. into the abdominal skin around but not within 5 cm of the navel. If the abdomen appeared inappropriate according to the Investigator's judgment, the study drug was injected s.c. into the thigh.

Hydroxyurea
HU was available as 500 mg capsules, containing 500 mg of hydroxycarbamide; the capsule could also have contained (b) (4) citric acid, (b) (4) gelatin, (b) (4), lactose, magnesium stearate, (b) (4), sodium monohydrogen phosphate, (b) (4) and titanium dioxide.

Randomization Schema:
This was a randomized study with two treatment arms (AOP2014 and HU). Treatments were assigned to individual patients based on stratified randomization in order to ensure random assignment to treatments and balanced distribution of both treatments within each stratum. There were three stratification factors and eight strata:

- Previous HU exposure (Yes/No)
- Age at screening (≤ 60 or > 60 years), and
- Presence of thromboembolic events in the past (Yes/no)

Eligible patients were randomized 1:1 to either AOP2014 or HU.

Primary Efficacy Endpoint:
The disease response rate after 12 months was defined by the number of patients meeting all four parameters:

- Hematocrit (Hct) < 45% without phlebotomy (at least 3 months since last phlebotomy)
- Platelets < 400 x 10^9/L
- Leukocytes < 10 x 10^9/L
- Normal spleen size (<12cm for females and <13 cm for males)

1.5.1.2 Statistical Methodologies

Sample Size Calculation.
Final analysis of this study included 12-month response data. One formal IA was carried out on 6-month response data.

PROUD-PV was originally designed as a superiority study; hence, the sample size was calculated based on a superiority comparison of AOP2014 vs HU with the anticipated size of treatment effect in an overall rate of responders at 12 months was 25% at least. The assumed rate of responders in patients treated with HU was 15% and 40% in patients treated with AOP2014. Supposed drop-out rate was approximately 20%. The dropped-out patients were considered as non-responders according the definition of the primary endpoint which would decrease assumed

rate of responders to 12% and 32%. Based on these assumptions, one hundred and twenty-six patients per group (252 in total) would be needed in order to detect the difference in response rate between treatment groups at 1% (two-sided) significance level with 90% power using standard chi-square test. Taking into account 8 strata, 128 patients per treatment group (256 in total) were planned to be enrolled.

Later, the sponsor proposed a change of the study primary objective from a superiority comparison to a non-inferiority comparison of AOP2014 to HU with a NI margin of 10.5% in the study protocol version 4.0 dated 15 June 2016, the statistical power for non-inferiority hypothesis was retrospectively calculated.

The last subject completed on 8 April 2016. Retrospective power calculations do not influence the enrollment. The Sponsor was still blinded to results of the study at time of power calculations. The power of the primary analysis was estimated for some scenarios, based on the blinded part ('open' report) of the DMC report prepared in April 2016 (file 'AOP Orphan PROUD-PV Report for Data Monitoring Committee OPEN VERSION 0.2 04APR2016.pdf'). The open report included descriptive statistics of both arms pooled.

Overall response rates and sample sizes (blinded analysis of pooled treatment arms) in the 8 strata were:

Stratum	N	Response rate*
S1: Age≤60, HU+, TE+	9	0%
S2: Age>60, HU+, TE+	9	13%
S3: Age≤60, HU- , TE+	11	40%
S4: Age>60, HU- , TE+	18	27%
S5: Age≤60, HU- , TE-	65	24%
S6: Age>60, HU- , TE-	61	26%
S7: Age≤60, HU+, TE-	43	16%
S8: Age>60, HU+, TE-	32	17%

*Drop-outs were distributed approximately uniformly across strata, and they were included as treatment failures in the estimation of response rates.
Source: Section 8.8 of SAP

Sponsor performed simulation studies based on the above sample sizes and overall response rates, generating 10, 000 studies for each scenario. Response rates were generated assuming a binomial distribution for each stratum, and equal treatment arm sizes within each stratum. The primary analysis was then performed for each simulated study, and the number of cases where the null hypothesis was successfully rejected were counted.

Results of the power calculation to conclude non-inferiority based on several NI margin are presented in the following table. Given a NI margin of -10%, the statistical power to conclude non-inferiority is roughly 4%.

Table 1: Power Calculation to Conclude non-inferiority

Difference between the response rates (test – reference)	Power to conclude non-inferiority
-0.1	4%
-0.08	8%
-0.06	15%
-0.04	25%
0	55%
0.04	81%
0.06	89%
0.08	95%
0.1	98%

Source: Section 8.8 of SAP

Efficacy Analysis Set*:*
Intent-to-Treat (ITT): all randomized patients.

Full Analysis Set (FAS): the intent-to-treat (ITT) population included all randomized patients, according to the treatment to which they were randomized, and irrespective of any protocol violations, subsequent therapies taken, etc. The FAS included the ITT population but excluded any patients for whom it was documented that they had taken no study medication and provided no follow-up (post-randomization) data.

Per Protocol Set (PPS): PPS consisted of patients included in the FAS who completed a certain pre-specified minimal exposure to the treatment regimen, had all measurements needed for assessment of the primary endpoint available and did not violate the study protocol in major concerns. It is used for efficacy sensitivity analysis.

Statistical Analysis of Primary Efficacy Endpoint:

Statistical Hypothesis:

$$H_0: p_A \leq p_H - 10.5\%$$
$$H_{a:}\ p_A > p_H - 10.5\%$$

Where p_A is the proportion of the responders in AOP2014 arm and p_H is the proportion of responders in HU arm and -10.5% is the NI margin.

The null hypothesis is the proportion of responders treated with AOP2014 is inferior by at least 10.5% to the proportion of responders treated with HU after 12 months. The alternative hypothesis is the non-inferiority of AOP2014 with a non-inferiority margin of 10.5%. The one-sided null hypothesis was tested against the alternative hypothesis at a one-sided significance level of 2.5%. This was tested through estimating the two-sided 95% CI of $p_A - p_H$ and comparing the lower limit of the CI to the non-inferiority margin of -10.5%.

Primary Analysis:

The primary analysis was conducted using a weighted Cochran-Mantel-Haenszel test (using SAS® terminology). Corresponding response rate difference between the two treatment arms (test reference) and its 95% confidence interval were calculated. The stratification factors were those used in the randomization scheme, namely: age groups (≤60 years vs. >60 years), previous thrombotic event (presence vs absence), and previous HU exposure (yes vs. no).

Non-inferiority was concluded in the lower limit of the 95% two-sided CI of the Mantel-Haenszel common estimate of response rate difference exceeds -0.1050 (-10.5%).

For the primary endpoint, failing to meet any one of the criteria for a response was sufficient to identify a patient as non-responding, even if other parts of the assessment were missing. Therefore, patients who completed the study but have missing components to the assessment of response could still be able to be properly included in the analysis. Patients, who withdraw early for any reason, were classified as treatment failures.

The same drop-out rate in both treatment groups had been supposed in sample size calculation; therefore, violation of this assumption was checked before database (DB) lock. Because proportion of randomized patients who dropped-out during the study was balanced between the two treatment arms, no adjustment in statistical analysis was needed.

Multiplicity Adjustment:
No statistical adjustments were made for multiple comparisons. Only primary efficacy endpoint, disease response rate was assessed in this review.

Interim Analysis:
The DMC was responsible for monitoring accumulating data. Based on recommendation of the DMC, the formal interim analysis was not performed. The results of comparison of treatment arm using the 6-month data were only provided to the DMC as a part of 'Closed' DMC report for regular DMC meeting. The sponsor received only the 'Open' report where descriptive statistics of both treatment arms pooled were presented. Therefore, the sponsor was blinded against the results of the study until the final analysis. According to the DMC statement no sample size re-assessment was needed and therefore not performed.

Reviewer's Comment: *There are several issues with the design and analysis of PROUD-PV study. The sponsor original designed PROUD-PV to demonstrate a superiority of AOP2014 compared to HU. In the amended clinical study protocol version 4.0, sponsor changed the superiority comparison to a non-inferiority (NI) comparison of AOP2014 to HU with a NI margin of 10.5% and a retrospectively calculated statistical power of 4%. A study with a statistical power of 4% is not considered a well-designed study. Further, the change from a superiority comparison to a non-inferiority comparison was made ~2 months after the completion of the study (08 April 2016). Given that the study was open-label and the sponsor had access to the unblinded data during the study, the modification to the study design and analysis method based on data may violate the integrity of the study and the results may not be reliable or interpretable. In addition, the NI margin of 10.5% was not justified. In a later meeting, the sponsor further proposed to change the NI margin from* (b) (4) *At the time*

of these proposals, the review team did not agree to these changes. These correspondences can be viewed in the meeting minutes dated 20 April 2017, 11 April 2019, and 04 September 2019.

1.5.1.3 Patient Disposition, Demographic and Baseline Characteristics

Patient Disposition

Two hundred and fifty-seven patients were randomized (Table 2). Three patients in the HU arm never received the treatment drug. Hence, there were 254 patients in the full analysis set or (FAS). FAS was defined as randomized patients without those with no study medication or no post-randomization data. Sponsor used FAS as the primary analysis population. Two hundred and twenty-nine patients were in the per protocol set (PPS). PPS was defined as patients included in the FAS who completed a certain pre-specified minimal exposure to the treatment regimen, had all measurements needed for assessment of the primary endpoint available, and did not violate the study protocol in major concerns. It is used as sensitivity analysis.

Two hundred and seventeen patients completed the study. Twenty-one (16.5%) randomized patients in AOP2014 arm and 16 (12.3%) of randomized patients in HU arm discontinued early. The main reason for early discontinuation was AE related events or withdrawal by patients.

Table 2: Disposition of Study Patients (PROUD-PV)

	AOP2014 N (%)	HU N (%)	Total N (%)
Randomized (ITT)	127 (100.0%)	130 (100.0%)	257 (100.0%)
Treated	127 (100.0%)	127 (97.7%)	254 (98.8%)
Full Analysis Set (FAS)	127 (100.0%)	127 (97.7%)	254 (98.8%)
Per Protocol Set (PPS)	115 (90.6%)	114 (87.7%)	229 (89.1%)
Completed	106 (83.5%)	111 (85.4%)	217 (84.4%)
Discontinued	21 (16.5%)	16 (12.3%)	37 (14.4%)
Reason for Early Discontinuation			
Administrative reasons (e.g. moving, prior commitments)	4 (3.1%)	1 (0.8%)	5 (1.9%)
Adverse events (AEs)	5 (3.9%)	2 (1.5%)	7 (2.7%)
Following AEs, independent of the recovery/outcome status*	6 (4.7%)	1 (0.8%)	7 (2.7%)
Lack of efficacy	0	2 (1.5%)	2 (0.8%)
Other	0	5 (3.8%)	5 (1.9%)
Withdrawal of consent	6 (4.7%)	5 (3.8%)	11 (4.3%)

* Following adverse events, independent of the recovery/ outcome status: symptomatic autoimmunity, HADS score of 11 or higher on either or both of the subscales, suicidal ideation or any other psychiatric drug-related adverse events of moderate or severe intensity, pulmonary, ocular or cardiac drug-related toxicity of moderate or severe intensity.
Source: FDA analysis

Baseline Demographic and Disease Characteristics:
Patients' Baseline Demographic Characteristics data are shown in Table 3. Baseline demographic characteristics were balanced between the AOP2014 arm and HU arm. The median age was 60 years (range 21 to 85 years). 53.5% and 52.8% of patients in AOP2014 arm and HU arm respectively were female. 100% of patients were White. All participants were from 13 countries in Europe and were divided into Western or Eastern Regions. 84.3% of these participants were from Eastern Europe. Most participants (63.0%) were not previously treated with HU. The median duration of PV was 1.9 months (range 0-146 months) in the AOP2014 treatment arm and 3.6 months (range 0-126 months) in the HU treatment arm indicating that patients were diagnosed in an early stage of the disease.

Table 3: Baseline Demographic Characteristics – FAS (PROUD-PV)

	AOP2014 (N=127)	HU (N=127)
Age, years		
N	127	127
Mean (±SD)	58.5 (±10.81)	57.9 (±13.10)
Median	60.0	60.0
Min, Max	30, 85	21, 81
Gender, [n (%)]		
Female	68 (53.5)	67 (52.8)
Male	59 (46.5)	60 (47.2)
Race, [n (%)]		
White	127 (100.0)	127 (100.0)
Region, [n (%)]		
Western Europe	20 (15.7)	20 (15.7)
Eastern Europe	107 (84.3)	107 (84.3)
Previous treatment with HU, [n (%)]		
Yes	47 (37.0)	47 (37.0)
No	80 (63.0)	80 (63.0)
Duration of PV (month)		
N	127	127
Mean (±SD)	12.6 (±24.70)	15.7 (±25.65)
Median	1.9	3.6
Min, Max	0, 146	0, 126

Source: FDA analysis

As noted, PROUD-PV was conducted in 13 countries in Europe. Most patients were from countries in Eastern Europe (84.3%). The distribution of patients by country is show below.

Table 4: Baseline Information by Country – FAS (PROUD-PV)

Country, [n (%)]	AOP2014 (N=127)	HU (N=127)	Total (N=254)
Western Europe	20 (15.7)	20 (15.7)	40 (15.7)

Spain	1 (0.8)	0	1 (0.4)
Italy	0	1 (0.8)	1 (0.4)
France	5 (3.9)	8 (6.3)	13 (5.1)
Germany	3 (2.4)	3 (2.4)	6 (2.4)
Austria	11 (8.7)	8 (6.3)	19 (7.5)
Eastern Europe	107 (84.3)	107 (84.3)	214 (84.3)
Bulgaria	25 (19.7)	20 (15.7)	45 (17.7)
Czech Republic	14 (11.0)	12 (9.4)	26 (10.2)
Hungary	14 (11.0)	20 (15.7)	34 (13.4)
Poland	12 (9.4)	15 (11.8)	27 (10.6)
Romania	3 (2.4)	6 (4.7)	9 (3.5)
Russian Federation	16 (12.6)	17 (13.4)	33 (13.0)
Slovakia	6 (4.7)	4 (3.1)	10 (3.9)
Ukraine	17 (13.4)	13 (10.2)	30 (11.8)

Source: FDA analysis

At baseline, the hematological parameters, spleen size, and JAK2 allelic burden were also comparable between the two treatment arms

Table 5: Baseline Disease Characteristics – FAS (PROUD-PV)

	AOP2014 (N=127))	HU (N=127)
Hematocrit		
N	124	126
Mean (±SD)	49.5 (±5.43)	49.8 (±5.49)
Median	49.6	49.5
Min, Max	37.7, 65.8	38.5, 70.8
Leukocytes		
N	124	125
Mean (±SD)	12.2 (±5.29)	12.6 (±5.18)
Median	10.9	10.9
Min, Max	4.6, 32.9	4.9, 30.1
Platelets		
N	124	126
Mean (±SD)	556.7 (±256.96)	528.3 (±236.43)
Median	506.5	466.5
Min, Max	149, 1065	135, 1065
Normal Spleen Size (spleen size ≤ 12 cm for females, ≤ 13 cm for males)		
Yes	59 (46.5%)	59 (46.5%)
No	68 (53.5%)	68 (53.5%)
JAK2V617F allelic burden (%)		
N	126	125
Mean (±SD)	41.9 (±23.49)	42.8 (±24.14)

Median	37.3	37.4
Min, Max	0, 95	0, 90

Source: FDA analysis

1.5.1.4 Efficacy Results

Primary Endpoint:
Total of 217 patients completed the study in the FAS population, 37 patients discontinued the study were considered as non-responders. Further, A total of 9 patients in FAS had no data for disease response criteria at 12-month visit available. Four patients in each treatment group had no data for disease response criteria available but other criteria for disease response were fulfilled. For one patient in the AOP2014 treatment arm, no data for disease response evaluation was available at all. Sponsor excluded these patients from the FAS population, hence, there are 245 patients in the sponsor defined FAS population (122 patients in AOP2014 and 123 patients in HU) for evaluation of the disease response. This reviewer considered these 9 patients also as non-responders and included in the FAS population for evaluation. The results from both definitions were similar and presented in table below.

Since the lower bound of 95% CI of the difference (AOP2014 – HU) in disease response was less than the NI margin of -10.5% for both FAS populations, the non-inferiority of AOP2014 to HU was not demonstrated at 12-months.

Table 6: Evaluation of Non-inferiority in Disease Response Rate (PROUD-PV)

	AOP2014	HU
FAS Population (FDA)		
N	127	127
Responders, n (%)	26 (20.5%)	34 (26.8%)
Difference (%) (AOP2014 - HU)	-6.3	
95% CI	-16.7, 4.1	
FAS Population (Sponsor)		
N	122	123
Responders, n (%)	26 (21.3%)	34 (27.6%)
Difference (%) (AOP2014 - HU)	-6.6	
95% CI	-17.2, 4.1	

Source: FDA analysis

Sensitivity Analyses:

Different Analysis Population:

In addition to the primary analysis population, response rate in intent-to-treat (ITT) and per-protocol set (PPS) populations were observed in AOP2014 and HU arms and the results are

shown in table below. Results from both ITT and PPS populations agree with the results from the FAS population. This supports the conclusion that non-inferiority of AOP2014 to HU was not demonstrated with the 12-months data.

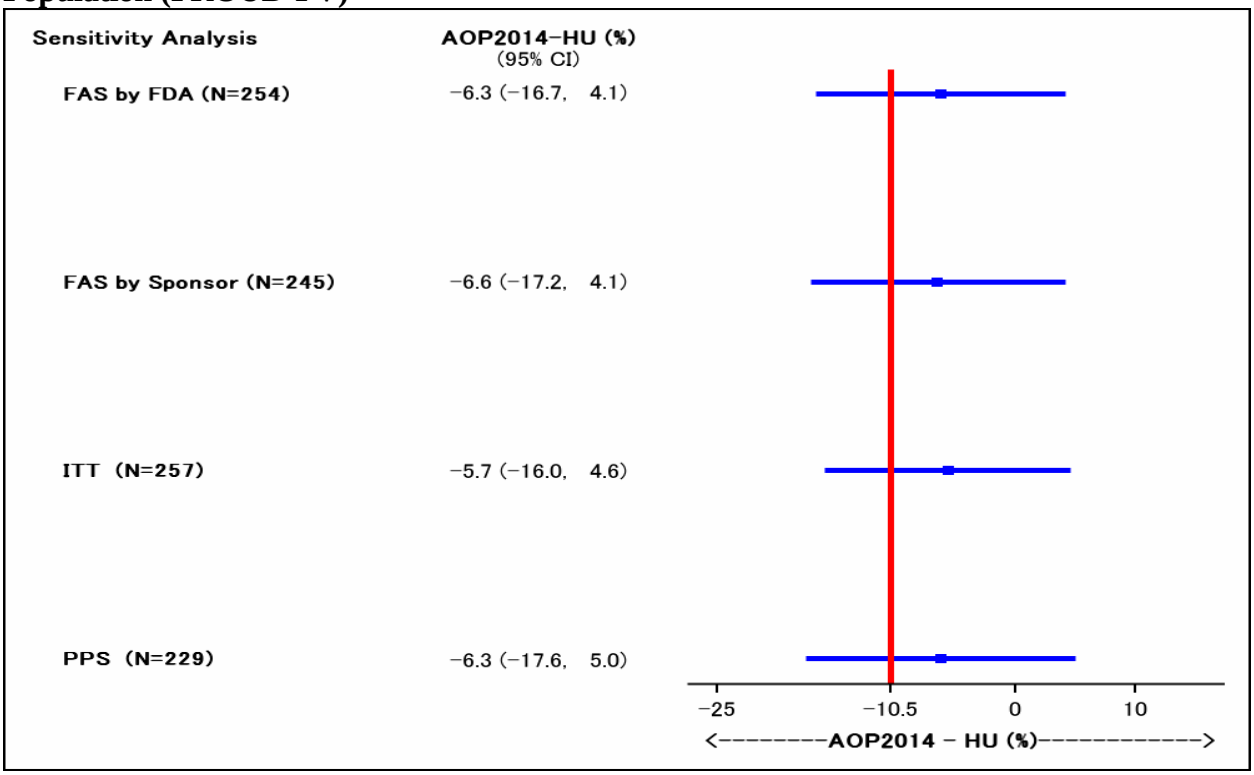
Table 7: Sensitivity Analysis for the Disease Response Rate with Different Analysis Population (PROUD-PV)

	AOP2014	HU
ITT Population		
N	127	130
Responders, n (%)	26 (20.5%)	34 (26.2%)
Difference (%) (AOP2014 - HU)	-5.7	
95% CI	-16.0, 4.6	
PPS Population		
N	115	114
Responders, n (%)	26 (22.6%)	33 (28.9%)
Difference (%) (AOP2014 - HU)	-6.3	
95% CI	-17.6, 5.0	

Source: FDA analysis

A graphical illustration of the results from the different analysis population is presented in the forest plot below.

Figure 2: Forest Plot for Sensitivity Analysis of Primary Endpoint with Different Analysis Population (PROUD-PV)



Source: FDA analysis

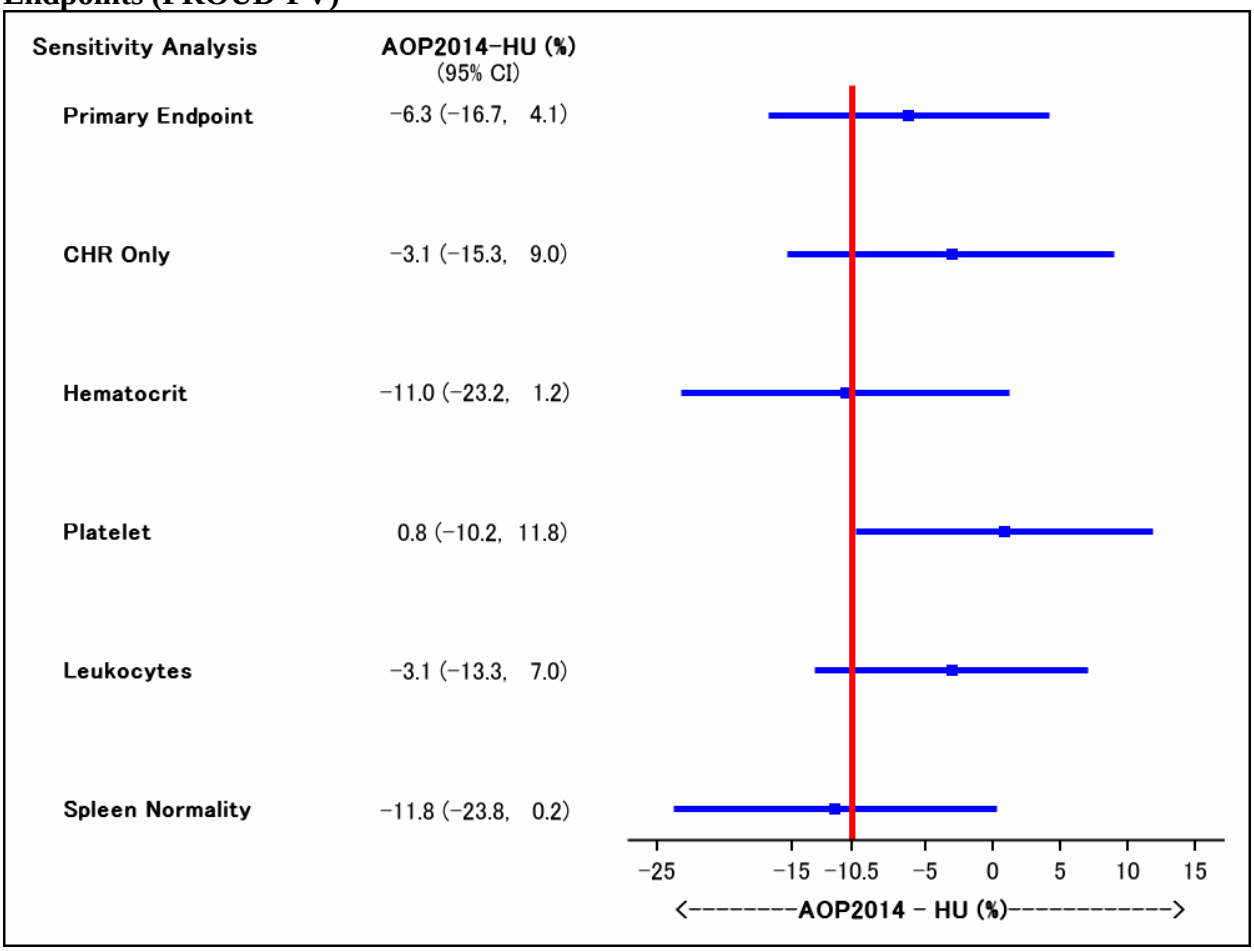
Exploratory Endpoints:

Complete Hematological Response (CHR) at Month 12 was defined as follow:

1. Hematocrit < 45% without phlebotomy (at least 3 months since last phlebotomy)
2. Platelets < 400 x 10^9/L
3. Leukocytes < 10 x 10^9/L

Complete hematological response (CHR) and each individual component of the primary endpoint were analyzed as sensitivity analyses (Figure below). These results also agree with the result from the primary endpoint.

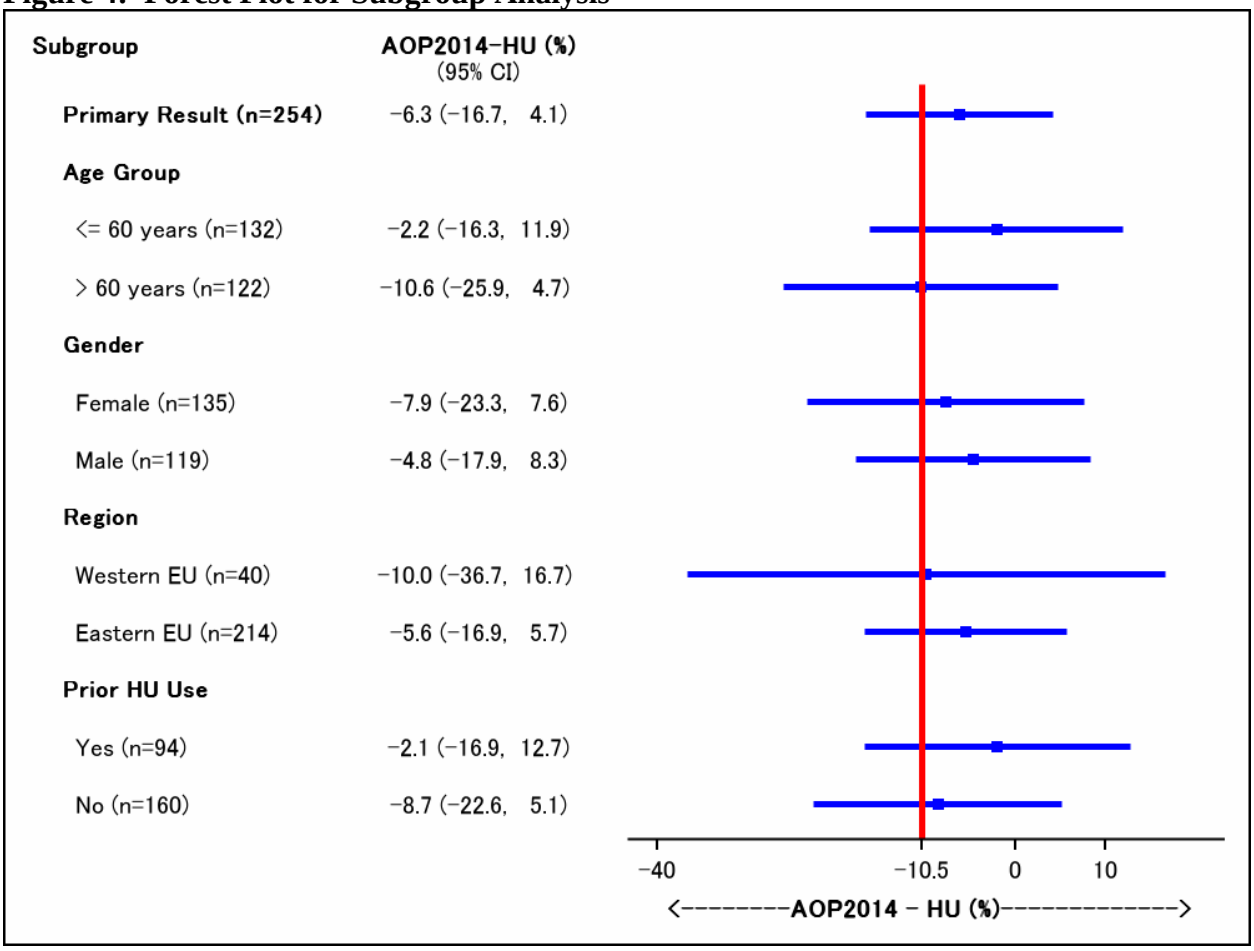
Figure 3: Forest Plot for Sensitivity Analysis of Primary Endpoint with Exploratory Endpoints (PROUD-PV)



Subgroup Analysis:

Subgroup analysis of the primary endpoint by age group, sex, region, and prior HU use were assessed and the results agree with the primary analysis result. No outlier subgroups were observed (Figure below).

Figure 4: Forest Plot for Subgroup Analysis

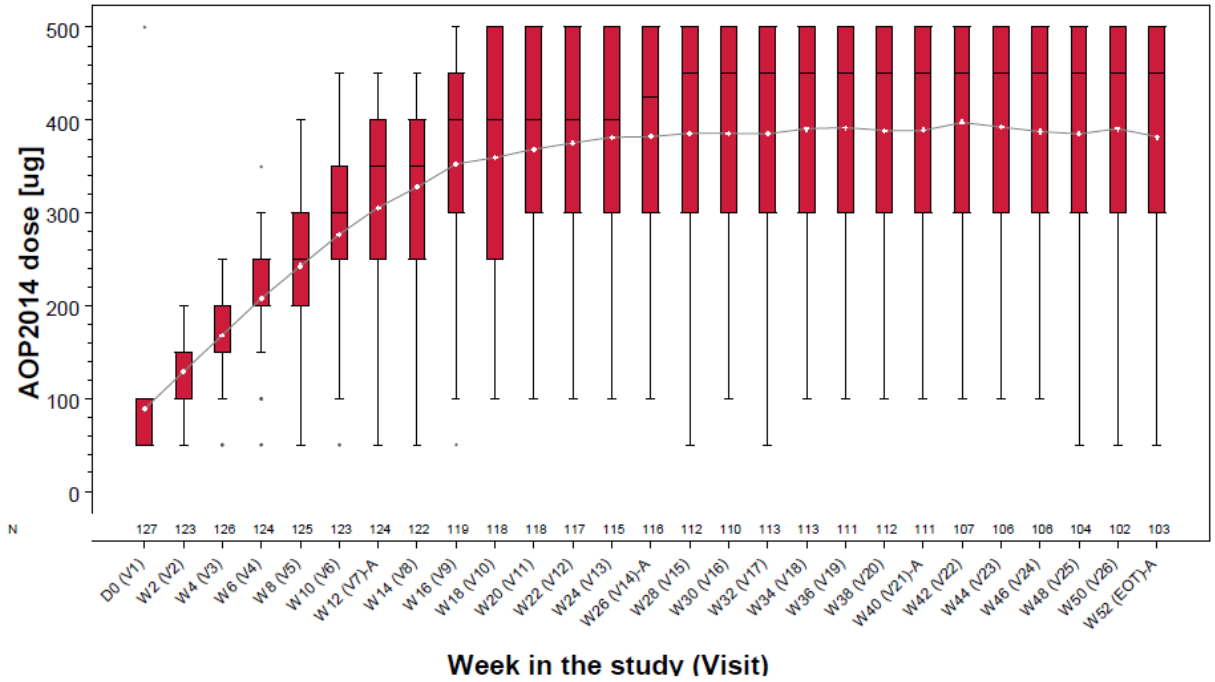


1.5.1.5 Possible Reason for Study Failure

A possible reason for the trial failure includes the different dose titration between the two arms. The two arms in the study had different dose titrations with a very conservative dose titration for AOP2014. The HU arm started at a dose of 500mg and was titrated upward to 3000mg with a dose plateau achieved at 8 weeks. The dose escalation in the AOP2014arm was intentionally done slowly with the goal to limit or minimize toxicity. The AOP2014 arm started at 100ug and the dose increased in 50ug steps up to 500ug every 2 weeks. The time to reach maximum dose was not achieved until week 28 week. There was a 20-week delay (28 weeks versus 8 weeks in the HU arm) in the AOP2014 arm to reach maximum dose. Dose reductions for dose related toxicity were twice as high in the HU arm. The PROUD-PV study duration was too short, and the sponsor's assessment of efficacy was planned too early at 12 months.

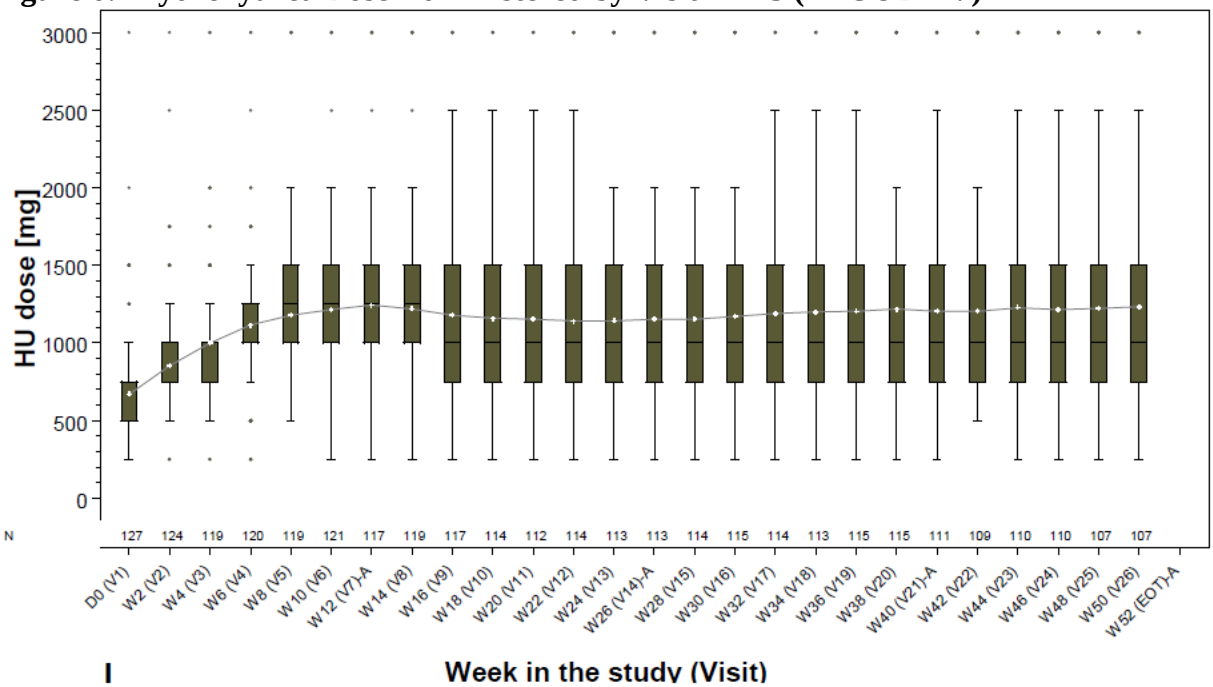
The following figures display the AOP2014 and HU by dose on the y-axis and study date on the x-axis.

Figure 5: AOP2014 Dose Administered by Visit -FAS (PROUD-PV)



Source: Figure12.1-1 of Sponsor's CSR for PROUD-PV

Figure 6: Hydroxyurea Dose Administered by Visit - FAS (PROUD-PV)



Source: Figure12.1-2 of Sponsor's CSR for PROUD-PV

Thus, the review team reviewed the results from the PEGINVERA (phase I/II study) and the single-arm experience efficacy data at 24 and 36 months from the PROUD-PV/ CONTINUAITON study.

1.5.2 CONTINUATION-PV (Exploratory Study)

1.5.2.1 Study Design and Endpoints

Study CONTIUATION-PV was a phase IIIb, open-label continuation of the PROUD-PV study performed in adults diagnosed with PV. It was designed as an exploratory study to assess the long-term efficacy of AOP2014 (P1101) or best available therapy (BAT).

Patients who received AOP2014 or HU in the primary study, PROUD-PV Study, and who completed the respective study, were eligible in the CONTINUATION-PV Study. After having signed the informed consent form (ICF), eligible patients were to be enrolled. Patients who received AOP2014 in the PROUD-PV Study were to be assigned to the AOP2014 arm and HU patients were to be allocated to the BAT arm.

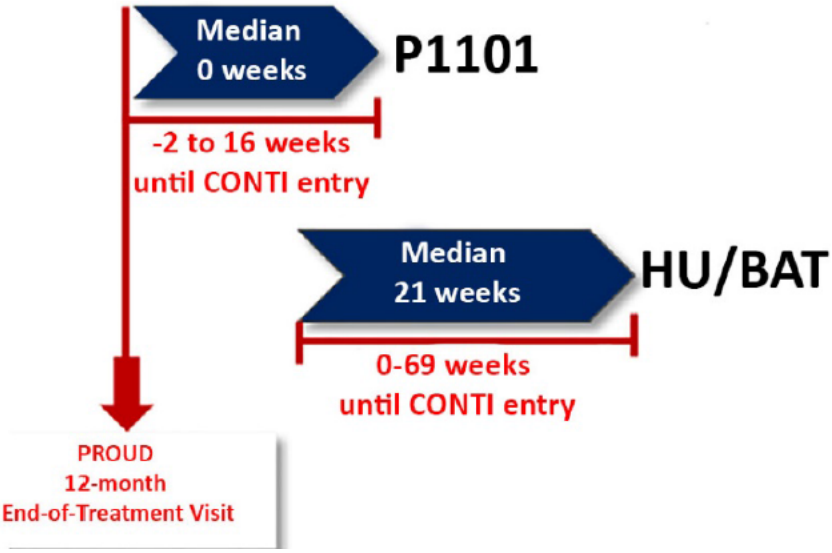
Primary Efficacy Objective:

- To assess the long-term efficacy of AOP2014 or standard first line treatment (HU or other best available treatment [BAT], referred to as 'Control') in terms of disease response rate in patients diagnosed with PV, who were previously participated in the PROUD-PV Study and who completed this study.

- To assess the long-term efficacy including changes in disease burden in patients diagnosed with PV. Disease burden is defined as disease-related signs (clinically significant splenomegaly) and disease-related symptoms (microvascular disturbances, pruritus, headache), assessed by investigator.

CONTIUATION-PV was originally designed as a single-arm extension study only for patients receiving AOP2014 from Study PROUD-PV, and the BAT group was added later. The first AOP2014 patient in the CONTINUATION-PV study was enrolled on 25 NOV 2014. The Amended Protocol Version 5.1, which added the HU/BAT group, was dated Nov 2016. The first patient being enrolled in the HU/BAT group completed the PROUD study on 19 OCT 2015 and enrolled in the CONTINUATION-PV study on the same day. This resulted in a gap of ~11 months between the starting date of the AOP2014 group and that of the HU/BAT group, the mean time between the end of PROUD-PV and the start of CONTINUATION-PV was 1 week for patients in the P1101 group and 26 weeks for the HU/BAT group.

Figure 7: Study Entry Gap Between AOP2014(P1101) and BAT Participants After PROUD-PV (CONTINUATION-PV)



Source: Figure 4 of sponsor submitted integrated-lta-report

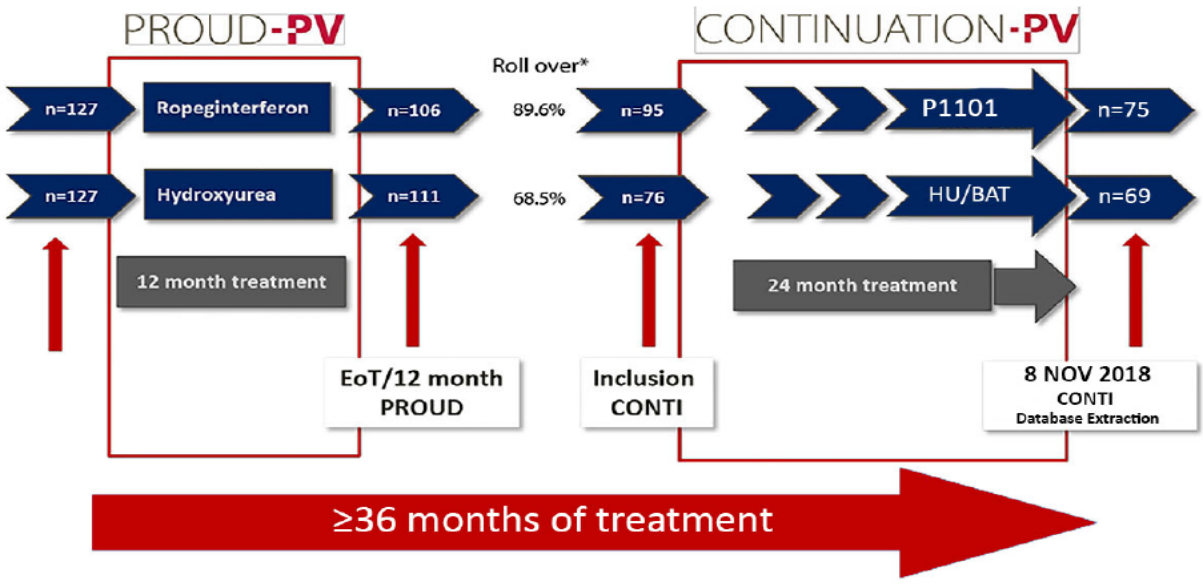
Table 8: Time from End-of-Treatment of the PROUD-PV Study to Informed Consent Date of the CONTINUATION-PV Study (Weeks)

	AOP2014	HU	ALL
Enrolled Subjects	95	76	171
Mean (SD)	1.1 (3.35)	26.3 (21.73)	12.3 (19.30)
Median	0.0	21.1	0.0
Min, Max	-2.0. 16.0	0.0, 68.7	-2.0, 68.7

Source: Table 2 of sponsor submitted integrated-lta-report

By the time the HU/BAT group began enrollment, some patients from the HU group had already completed the PROUD-PV study and had been treated by their own physicians for periods up to a year or more. Therefore, the delay in adding the HU/BAT group could have resulted in a lower entry rate from the PROUD-PV study to the CONTINUATION-PV study for the HU/BAT group compared with the AOP2014 group.

Figure 8: Rollover Rates from PROUD-PV to CONTINUATION-PV



Source: Figure 6 of sponsor submitted integrated-lta-report

Selection of Study Population:
Patients were eligible to enroll into the CONTINUATION-PV from PROUD-PV if they met all of the following inclusion criteria and none of the exclusion criteria.

Inclusion Criteria:

1. Patients who had completed the 12 months PROUD-PV Study and who fulfilled at least one of the following criteria:

    a) Normalization of at least two out of three main blood parameters (Hct, PLTs and WBCs) if these parameters were moderately increased (Hct<50%, WBC<20 x 109/L, PLTs<600 x 109/L) at baseline of the PROUD-PV Study, OR

    b) >35% decrease of at least two out of three main blood parameters (Hct, PLTs and WBCs) if these parameters were massively increased (Hct>50%, WBCs>20 x 109/L, PLTs >600 x 109/L), at baseline of the PROUD-PV Study, OR

    c) Normalization of spleen size, if spleen was enlarged at baseline of the PROUD-PV Study, OR

    d) Otherwise a clear, medically verified benefit from treatment (e.g. normalization of disease-related micro-vasculatory symptoms, substantial decrease of JAK2 allelic burden).

2. Signed written ICF.

Exclusion Criteria:

1. Non-recovery from the AOP2014 related toxicities to the grade (usually, grade I) which allowed continuation of the treatment.

2. HADS score of 11 or higher on either or both of the subscales, and /or development or worsening of the clinically significant depression or suicidal thoughts.

3. Progressive and clinically significant increase of liver enzyme levels despite dose reduction, or if such increase was accompanied by increased bilirubin level or any signs or symptoms of a clinically significant autoimmune disease.

4. Clinically significant development of a new ophthalmologic disorder, or worsening of a preexisting one, during the study.

5. AOP2014 arm only: Loss of efficacy of AOP2014 or any comparable situation where no further benefits of treatment continuation were expected by the Investigator.

Reviewer's Comment: *CONTINUATION-PV was an extension of PROUD-PV (randomized, active- controlled study). According to sponsor's response to FDA Information Request (IR) on 03 September 2019, patients who were eligible to enroll into CONTINUATION-PV were a subgroup of PROUD-PV that would need to show some measure of benefit (refer to inclusion criteria). This selection criteria could introduce bias. If the patient population in CONTINUATION-PV contains only patients who respond the treatment well, then these patients may be different from the patients studied in PROUD-PV. As a result, the estimated disease response rate from the study maybe "over-optimistic" and not a good estimation of the true disease response rate in the proposed disease population. In addition, patients were assigned to treatment arms (AOP2014 vs. BAT) not based on a randomization scheme but were based on the treatment arms they were assigned to in PROUD-PV study.* W*ithout proper randomization, selection bias may also be introduced, that could lead to the presence of confounders (either known or unknown) and undermine the study results and increase the type I error rate. Further, since BAT arm was added much later to the study, patients retention in the two treatment arms differed substantially between the two treatment arms (89.6% (95/106) patients who completed AOP2014 arm in PROUD-PV enrolled in the AOP2014 arm of CONTINUATION-PV and only 68.5% (76/111) patients who completed HU arm in PROUD-PV enrolled in the BAT arm of CONTINUATION-PV). This differential enrollment could also introduce bias.*

Co-Primary Efficacy Endpoints:

The primary efficacy endpoint is the rate of disease response at month 24 and 36. The two co-primary efficacy evaluation criteria are:

1. Disease response defined as:
    a. Hematologic response (CHR) (Hct<45% without phlebotomy [at least 3 months since the last phlebotomy]), PLTs<400 x 10^9/L and WBCs<10 x 10^9/L), and
    b. Normal spleen size.

2. Disease response defined as:
    a. Hematologic response (CHR) (Hct<45% without phlebotomy [at least 3 months since the last phlebotomy], PLTs<400 x 10^9/L and WBCs<10 x 10^9/L), and
    b. Resolution and/or clinically improvement of disease related signs (clinically significant splenomegaly) and disease-related symptoms (microvascular disturbances, pruritus, headache).

The patient was to be classified as "responder" (treatment success) only if all the disease response criteria were met. If any criterion was not met, the patient was to be classified as a "non-responder" (treatment failure), regardless whether data for evaluation of other criteria were available or not. Patients withdrawn for any reason were to be classified as treatment failures.

Normal spleen size is defined as ≤ 12 cm in females and ≤ 13 cm in males. A spleen length of > 12 cm in females and > 13 cm in males is defined as enlarged, however, other characteristics like body mass may be considered in spleen size assessment. Normality of spleen size as evaluated by the investigator was to be used for analysis of the CONTINUATION-PV Study.

1.5.2.2 Statistical Methodologies

Sample Size Calculation.
No formal hypothesis was planned to be tested in the CONTINUATION-PV Study; therefore, no power calculation or sample size calculation was performed at planning phase of the study. Out of 217 patients who completed the PROUD-PV Study, 171 patients (95 patients in AOP2014 arm and 76 patients in BAT arm) continued in CONTINUATION-PV Study. The expected drop-out rate in the CONTINUATION-PV Study is up to 20%.

Efficacy Analysis Set*:*
Enrolled patients will be included.

Statistical Analysis for the Primary Efficacy Endpoint:
Comparison of treatment arms were performed using log-binomial regression model implemented in SAS procedure PROC GENMOD with log link function, with specification of binomial distribution (link=log dist=binomial) and with repeated statement (“subject effect” is patient ID). Rate ratio (RR) of responders between treatment arms (AOP2014/Control) and its 95% CI were calculated from estimates of regression coefficients (RR=exp (estimate of regression coefficient). Outcome of the log-binomial model were presented using RR and its 95% CI and p-value. Effect of other covariates was presented by p-values and RR and its 95% CI or model estimates.

Reviewer’s Comment: *Since this study did not have a formal statistical hypothesis and no formal sample size and power calculation. All statistical analyses were to be performed for exploratory purposes. No valid statistical conclusion should be drawn from the results.*

1.5.2.3 Patient Disposition, Demographic and Baseline Characteristics

Disposition of patients:
Among 217 patients who completed the PROUD-PV Study, the roll-over rate into the CONTINUATION-PV Study was 78.8% (171/217) overall (89.6% [95/106] for AOP2014-treated patients and 68.5% [76/111] for HU-treated patients).

Table 9: disposition of both PROUD-PV and CONTINUTION-PV

	AOP2014 N (%)	BAT N (%)	Total N (%)
Patients participated in PROUD-PV (FAS)	127 (100.0)	127 (100.0)	254 (100.0)
Premature Discontinuation PROUD-PV Adverse Event Lack of Efficacy Lost to follow-up Withdrawal by Subject Other	21 (16.5) 11 (8.7) 0 (0.0) 4 (3.1) 6 (4.7) 0 (0.0)	16(12.6) 3 (2.4) 2 (1.6) 1 (0.8) 5 (3.9) 5 (3.9)	37 (14.6) 14 (5.5) 2 (0.8) 5 (2.0) 11 (4.3) 5(2.0)
Completed PROUD-PV(FAS)	106 (83.5)	111 (87.4)	217 (85.4)
Patients enrolled into CONTINUATION-PV	95/127 (74.8)	76/127(59.8)	171/254 (67.3)
Completed Proud and Enrolled in CONTINUATION-PV	95/106 (89.6)	76/111(31.5)	171/217 (78.8)
Premature discontinuation from CONTINUATION-PV Adverse Event Lack of Efficacy Lost to follow-up Withdrawal by Subject Other	20/95 (21.1) 7 (7.4) 3 (3.2) 3 (3.2) 3 (3.2) 4 (4.2)	7/76 (9.2) 2 (2.6) 0 (0.0) 0 (0.0) 3 (3.9) 2 (2.6)	27/171 (15.8) 9 (5.3) 3 (1.8) 3 (1.8) 6 (3.5) 6 (3.5)

Source: FDA analysis

Baseline Demographic and Disease Characteristics:
All patients enrolled in the CONTINUATION-PV study previously participated in the PORUD-PV study. The mean age at inclusion was 57.5 years with a median of 59 years (range 30 to 85 years). Eighty-three of 171 (48.5%) patients were male and 88/171 (51.5%) were female.

1.5.2.4 Efficacy Results

CO-Primary Endpoint:

Table 10 displays the sponsor's results for the primary endpoints and the statistical reviewer confirmed their results. The results for complete hematological response (CHR) and normal spleen size at Month 24 was response rate ratio of 1.1 (95% CI: 0.7, 1.7) (36.8% for AOP2014, and 30.3% for the BAT). At month 36, the response rate ratio was 1.4 (95% CI: 0.9, 2.1) (40.0% for AOP2014, 27.6% for the BAT).

The results for the CHR and improvement in disease burden at Month 24 demonstrate a response rate ratio of 1.4 (95% CI: 0.9, 1.9) (49.5% for AOP2014 and 34.2% for the BAT) and at Month 36, the response rate ratio was 1.4 (95% CI: 1.0, 2.0) (52.6% for AOP2014 and 36.8% for the BAT).

Table 10: Results for the Co-primary Endpoints (CONTINUATION-PV)

	AOP2014 (N=95)	BAT (N=76)
CHR and Normal Spleen Size		
Month 24 Responder, n (%)	35 (36.8)	23 (30.3)
Relative Risk (95% CI)	1.2 (0.7, 1.7)	
Month 36 Responder, n (%)	38 (40.0)	21 (27.6)
Relative Risk (95% CI)	1.4 (0.9, 2.1)	
CHR and Improvement in Disease Burden		
Month 24 Responder, n (%)	47 (49.5)	26 (34.2)
Relative Risk (95% CI)	1.4 (0.9, 1.9)	
Month 36 Responder, n (%)	50 (52.6)	28 (36.8)
Relative Risk (95% CI)	1.4 (1.0, 2.0)	

Source: FDA analysis

Reviewer's Comment: *Even though the comparative results between AOP2014 and BAT for the study seemed to show potential trend favoring AOP2014, no valid statistical conclusion can be drawn due to presence of potential bias from lack of randomization, lack of statistical power, lack of pre-specified hypotheses and success criteria, lack of proper statistical analysis, and lack of multiplicity control for multiple endpoints. Given these issues, the type I error rate will likely to inflated, hence, these results should be treated as exploratory only.*

1.5.3 SINGLE ARM INTERPRETATION OF AOP2014 (POST-HOC ANALYSIS)

1.5.3.1 Background

Given PROUD-PV was a failed study and the comparison between AOP2014 and BAT in CONTINUATION-PV was designed as exploratory only, the sponsor proposed to evaluate the long-term efficacy of AOP2014 based on the single arm experience at the mid-cycle meeting on 22 September 2020. This included all patients who were randomized into the AOP2014 arm of the PROUD-PV study and received at least one dose of AOP2014 (N=127). This sample size of 127 patients were considered "As-Treated" (AT) population and also can be considered as "Intent-to-Treat" population to the single arm of AOP2014. Sponsor further claimed that 127 patients is considered a sufficient sample size to evaluate response rate, response duration given the rarity of PV.

The sponsor further proposed to evaluate complete hematologic response (CHR) rather than the CHR with normal spleen size from PROUD-PV as the primary efficacy endpoint in this single-arm analysis because achievement of CHR, which includes control of hematocrit, white blood cell (WBC) and platelets, remains the core component of response assessment to cytoreductive therapy in PV, with confirmed predictive value for clinically relevant outcome. Further, spleen normality has limitations related to measurement of the spleen and the early stage of the disease population in the study. The PROUD-PV/CONTINUATION study was open-label and the components of the complete hematological response (laboratory measures) were objectively measured. The single arm data from PROUD-PV/CONTINUATION study at 24 and 36 months allow for assessment of responses beyond 12 months for CHR defined as HCT < 45% without phlebotomy (> 3 months since last phlebotomy), platelets, 400 x 10^9/L, leukocytes < 10 x 10^9/L. Using the population of 127, the evaluation of CHR was performed at month 12, 24 and 36. For patients who were enrolled in PROUD-PV but did not participate in the CONTINUATION study, they are treated as non-responders.

Primary Endpoints:

Complete Hematological Response (CHR) at Month 12:

- Hematocrit < 45% without phlebotomy (at least 3 months since last phlebotomy)
- Platelets < 400 x 10^9/L
- Leukocytes < 10 x 10^9/L

1.5.3.2 Efficacy Results:

The response rate for CHR and CHR plus normal spleen size at month 12, 24 and 36 with its 95% CI are show in table below.

Table 11: Results for the Response Rate - As Treated Population (Single Arm)

	PROUD-PV/CONTINUATION-PV (N=127)	
	Response Rate	95% CI

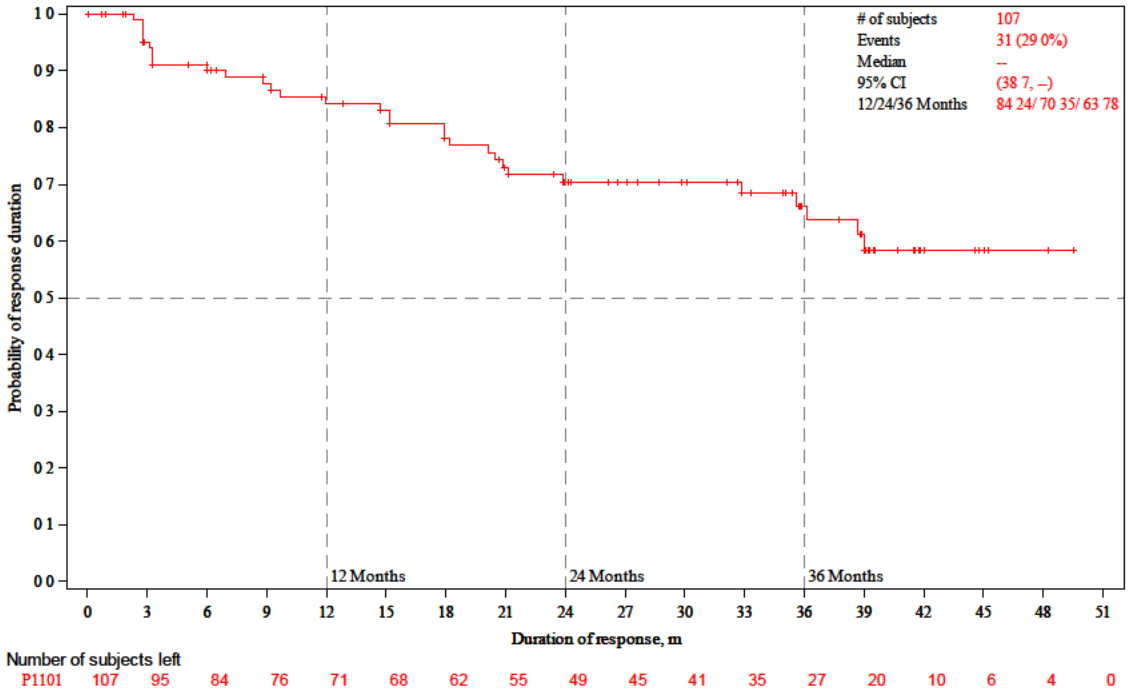
	N (%)	
CHR		
Month 12	66 (52.0)	(43.3, 60.7)
Month 24	67 (52.8)	(44.1, 61.4)
Month 36	67 (52.8)	(44.1, 61.4)
CHR plus Normal Spleen Size		
Month 12	31 (24.4)	(16.9, 31.9)
Month 24	34 (26.8)	(19.1, 34.5)
Month 36	38 (29.9)	(22.0, 37.9)

Source: Table 1 of Sponsor's FDA IR Response Final_20201015.docx

Duration of Response:

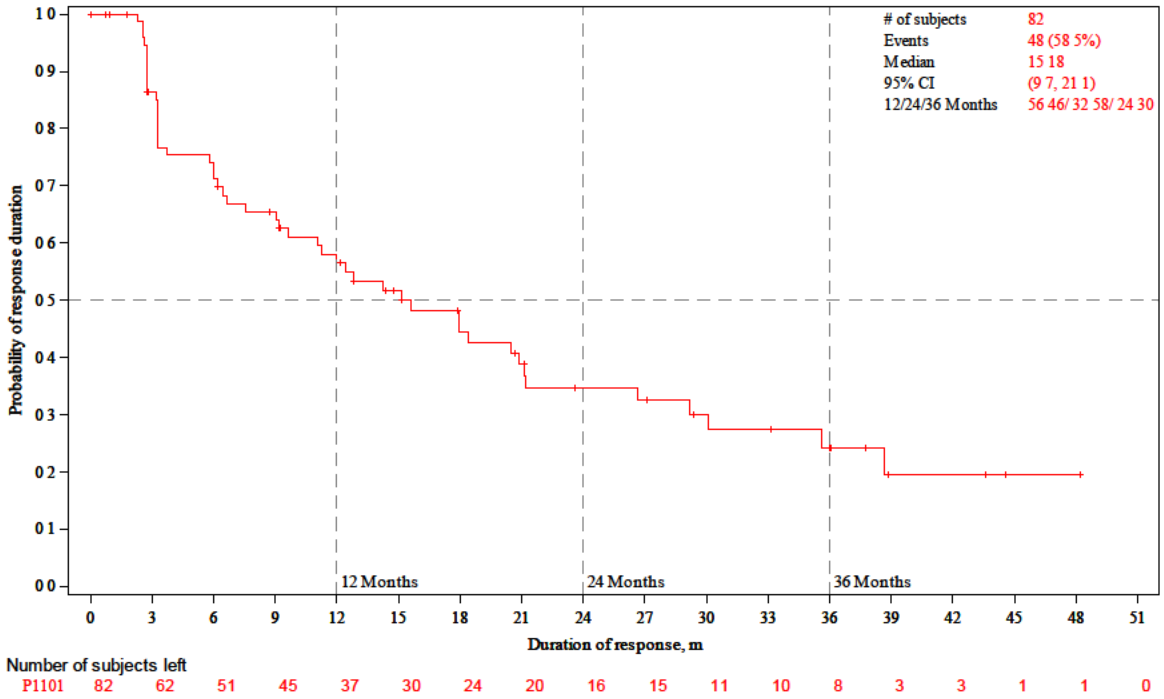
Duration of response for endpoints CHR and CHR plus normal spleen size are shown in the two figures below. For the endpoint CHR, the median duration of the 107 patients who reported responses was not reached. For the endpoint CHR plus normal spleen size, the median duration was 15.18 days with 95% CI (9.7, 21.1).

Figure 9: Kaplan-Meier Plot of Duration of Longest CHR - As Treated Population (Single Arm)



Source: Figure 1 of Sponsor's FDA IR Response Final_20201015.docx

Figure 10: Kaplan-Meier plot of Duration of Longest CHR with normal spleen size- As Treated Population (Single Arm)



Source: Figure 3 of Sponsor's FDA IR Response Final_20201015.docx

Table 12: Duration of Response (Single Arm)

Responders who attained response at:	N	>12 months	>24 months	>36 months
Month 12	66	95%	83%	78%
Month 24	67	95%	84%	78%
Month 36	67	97%	91%	84%

Source: FDA analysis

Regional Considerations:

The PROUD-PV/CONTINUATION study was conducted in Europe with majority of sites in Eastern Europe. The following table provides the hematology parameters at baseline for the different regions.

Table 13: Demographics by Region in PROUD-PV/CONTINUATION (Single Arm)

	All Countries N=127	Eastern Europe N=107	Western Europe N=20
Hematocrit, Median (min, max)	43 (36.8, 54.9)	43 (31, 50.9)	43.5 (38.6, 54.9)

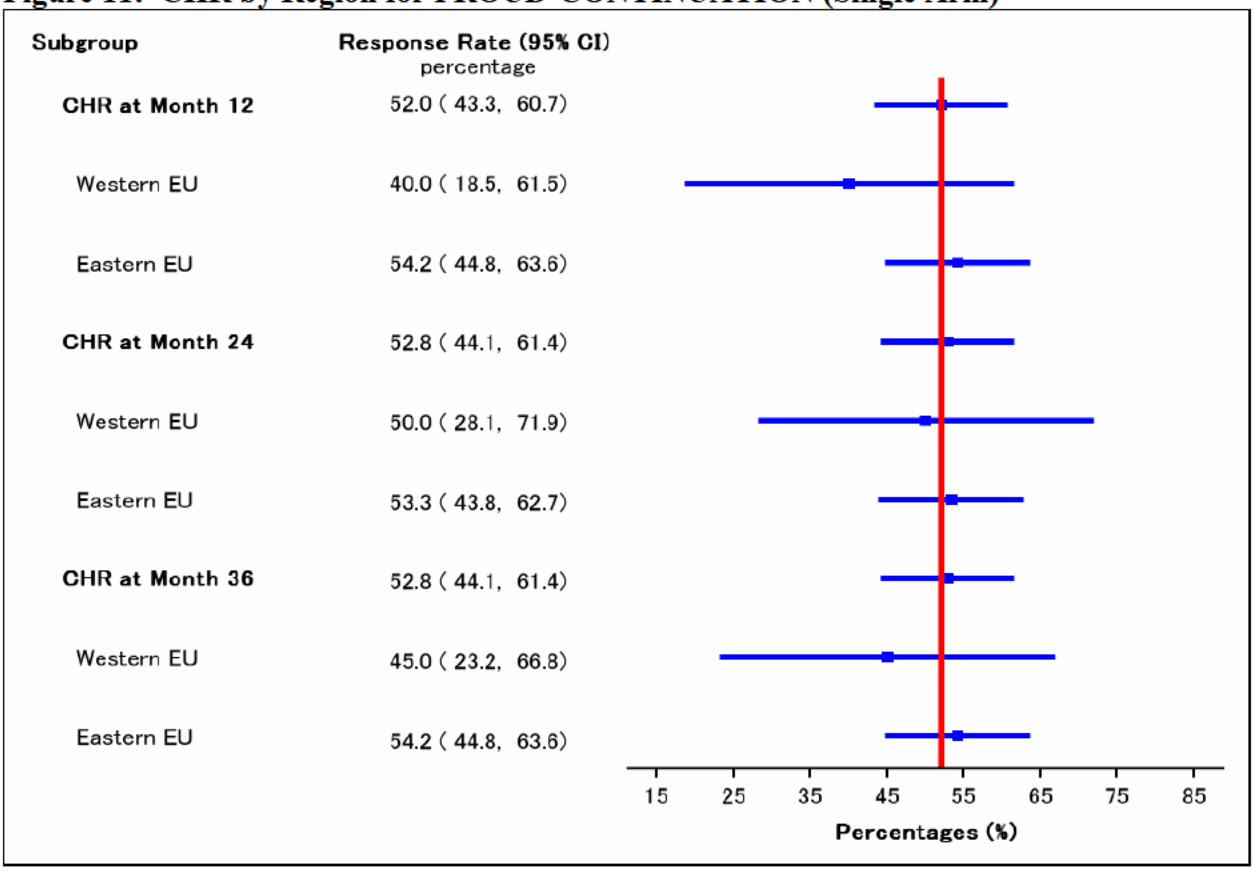
	All Countries N=127	Eastern Europe N=107	Western Europe N=20
Leukocyte, Median (min, max)	10.81(5.7, 25.6)	10.3 (4, 18.3)	13.1(7.8, 24.3)
Platelet, Median (min, max)	478 (138, 1293)	503 (138, 1293)	602.5 (69, 1242)

Source: FDA analysis

The baseline hematology parameters are similar between Western and Eastern Europe and applicability of results to US population was determined to be reasonable.

The following figure provides the CHR by Region

Figure 11: CHR by Region for PROUD-CONTINUATION (Single Arm)



Source: FDA analysis

1.5.4 PEGINVERA (PHASE I/II STUDY)

1.5.4.1 Study Design and Endpoints

The PEGINVERA study was an open-label, prospective, multicenter, phase I/II dose escalation study to determine the maximum tolerated dose and to assess the safety and efficacy of AOP2014, in patients with PV.

A diagnosis of PV was confirmed according to WHO criteria or Polycythemia Vera Study Group (PVSG) criteria plus *JAK2* mutation positivity including newly diagnosed, pre-treated and on cytoreductive therapy. Phase I was designed to determine the MTD of AOP2014using a 3+3 design in patients with PV. For the phase I portion, 40% (10/25) of the patients were undergoing treatment with HU. All 25 patients received one dose of study drug followed by a second dose 2 weeks later. The MTD was determined to be 540 mcg, the highest dose administered in the first treatment cycle as no DLTs (per protocol definition) or any other safety signals were observed for any of the dosing groups (dose range 50 to 540 mcg).

For the phase II part of the study, an additional 26 patients were enrolled. As in phase I, a diagnosis of PV was confirmed according to WHO criteria or PVSG criteria plus *JAK2* mutation positivity including newly diagnosed, pre-treated and on cytoreductive therapy. The 33.3% of the patients were undergoing treatment with HU and the time of enrollment.

Primary Endpoint:

- Efficacy Criteria was based on hematological responses defined as:
  - HCT < 45% (without phlebotomy in previous 2 months)
  - Platelet count ≤ 400 $x10^9$/L
  - WBC ≤10 x 10^9/L
  - Normal spleen size (longitudinal diameter ≤ 12 cm for females and ≤ 13 cm for males)
  - Absence of thromboembolic events

Patient Population:

The population in PEGINVERA included patients with newly diagnosed PV and patients who had received prior treatment for PV.

Inclusion criteria consisted of patients with confirmed diagnosis of PV according to the WHO criteria (2008) or the Polycythemia Vera Study Group criteria plus JAK-2 positivity and an ECOG PS less than or equal to 2.

Dose Titration:

The maximum dose tolerated (MTD) was determined to be 540ug and was the highest dose administered as no dose-limiting toxicities observed in Stage I of the study. In Stage II of the study, the starting dose of AOP2014 was 100ug with increases by 50ug every 2 weeks until a maximum dose of 500ug.

Sample Size Determination:

According to 3 + 3 escalation designs the sample size was conditioned by the design, i.e. there was a maximum number of 5 x 3 + 3 = 18 patients for determination of MTD in case 5 dose levels were to be investigated. An additional 6 to 12 patients were allowed following an ad-hoc decision. The total number of patients was however, no higher than 35. Stage 1 of the study was completed with 25 evaluable patients. A single cohort of 26 patients was added without formal size calculation after the MTD finding was completed to investigate further the drug efficacy and safety in PV.

1.5.4.2 Efficacy Result

Patient Disposition:

A total of 51 patients (25 patients from phase I and an additional 26 patients for phase II) enrolled from six center and twenty-five patients completed the study (13/25 in stage I and 12/26) in stage 2. Overall, 36 patients completed one year of treatment, 15 patients discontinued before week 50 and 11 discontinued after one-year treatment mainly due to treatment emergent adverse events. The Sponsor considered the full analysis set as 42 subjects (5 subjects excluded due to major protocol violations to include one patient with history of psychiatric disorder.

Baseline Demographic Characteristics:

The following table describes the demographics for the PEGINVERA study.

Table 14: Demographics of PEGINVERA Study

	PEGNIVERA N=51
Median (Range) age at trial entry	56 (35-82)
Male, [n (%)]	31 (60.8%)
Newly diagnosed PV, [n (%)] Prior Diagnosed PV, [n (%)]	8 (15.7%) 43 (84.3%)
Caucasian, [n (%)]	50 (98%)
Prior treatment with HU, [n (%)]	17(33%)
JAK2 Mutant Allele present baseline, [n (%)]	50(98%)
Hematological parameters, [mean (SD)] Leukocyte Hematocrit Platelet	 11.8 (5.2) 45.1 (4.0) 457.9 (186.5)

Source: FDA analysis

Efficacy Result:

The complete hematological response (normal spleen, no thrombosis and HCT < 45%, platelets ≤400 x 10^9/L, leukocyte count ≤10 x 10^9/L) was 27/51 (52.3%) (95% CI: 38.6, 66.0). The complete hematological response (CHR) with laboratory parameters only (HCT < 45%, platelets ≤400 x 10^9/L, leukocyte count ≤10 x 10^9/L) was 41/51 (80.4%) (95% CI: 66.9, 90.2). The median duration of CHR with laboratory parameters only was 20.8 months (range 13.0, 43.9).

Table 15: Duration of CHR in PEGINVERA Study

	Total (N = 51)
Subjects with response, n (%)	41 (80.4)
Events, n (%)	24 (58.5)
Censored, n (%)	17 (41.5)
Reason for censoring, n (%)	
No PD, No death	16 (39.0)
PD/Death after >=2 missed visits	1 (2.4)
Duration of longest response, months	
25th percentile (95% CI)	9.9 (2.2, 13.1)
Median (95% CI)	20.8 (13.0, 43.9)
75th percentile (95% CI)	NE (30.1, NE)
Min, max	0.0, 73.4

Source: FDA analysis

The Kaplan-Meier point estimate for 50% of patients to record a CHR was 1.4 years [521 days (95% CI 242, 992)]. The median time to response for CHR was 7.8 months.

Further evaluation of the individual components of CHR (leukocyte, HCT, and platelets) demonstrated a reduction from baseline for all three components.

- The mean baseline absolute leukocyte count was 11.5 (SD: 5.5) and by week 50 the mean leukocyte count was 5.8 (SD: 2.3).
- The mean baseline HCT was 45.1 (SD: 4.1) and by week 50 the mean HCT had decreased to 43.1 (SD: 3.9). Please note that the mean baseline HCT was close to 45 as patients were receiving phlebotomies prior to enrollment to help control HCT counts.
- The median baseline platelet count was 448.8 (SD: 182.0) and by week 50 had decreased to 234.7 (SD: 91.8).

1.6 Evaluation of Safety

Please refer to the clinical review on the safety issues for this study.

FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

1.7 Gender, Race, Age, and Geographic Region

Please refer to the previous sections.

1.8 Other Special/Subgroup Populations

We did not identify any other special populations.

1.9 Statistical Issues and Collective Evidence

The pivotal study, PROUD-PV, failed to demonstrate non-inferiority of study drug, AOP2014, to the active control, HU, in the primary endpoint, disease response rate. The CONTINUATION-PV study was designed as an extension study that offered only exploratory comparison between AOP2014 arm to BAT arm. After further evaluation, the review team decided that the efficacy evidence of AOP2014 can be based on the results from the PEGINVERA study and the results from the single-arm experience of PROUD-PV/CONTINUATION study. But PEGINEVERA is a single arm, phase I/II study and the results from the single arm experience from PROUND-PV/CONTINUATION is considered a post-hoc analysis. Hence, whether these results demonstrate substantial evidence of effectiveness should be based on clinical judgement considering unmet medical need.

1.10 Conclusions and Recommendations

Sponsor submitted two phase III studies (PROUD-PV and CONTINUATION-PV) and one phase I/II study (PEGINVERA) to show the effectiveness of AOP2014 compared to HU/BAT in the treatment of PV. After review, it was determined that PROUD-PV, the pivotal study, failed to show non-inferiority of AOP2014 to HU. In addition, the results from CONTINUATION-PV, an extension study of PROUD-PV, to compare the effectiveness of AOP2014 with BAT, were exploratory only and cannot be used to make any statistical inference.

Considering the unmet medical need for the treatment for PV and no spontaneous response for this disease, the review team evaluated the effectiveness of AOP2014 based on the results from PEGINVERA and the single arm experience in the combined PROUD-PV/CONTINUATION study. After evaluation, results from these two studies were determined to support the efficacy claim of AOP2014. This decision is mainly based on clinical judgement, not statistical inference.

1.11 Labeling Recommendations

Since PROUD-PV was a failed study and CONTINUATION-PV was an exploratory comparative study, the results from these two studies were not included in the label. (b) (4)

Reference:

Ayalew Tefferi, et al. (2014) 'Long-term survival and blast transformation in molecularly annotated essential thrombocythemia, Polycythemia Vera, and myelofibrosis'. *Blood (2014) 124(16): 2507-2513;* https://doi.org/10.1182/blood-2014-05-579136

Roberto Marchioli, et al. (2013) "Cardiovascular Events and Intensity of Treatment in Polycythemia Vera"**.** N Engl J Med 2013; 368:22-33 DOI: 10.1056/NEJMoa1208500

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

LOLA LUO
02/26/2021 06:38:28 PM

YEH FONG CHEN
02/26/2021 07:15:23 PM

THOMAS E GWISE
02/26/2021 08:36:16 PM