Statistical Analysis Plan

PharmaEssentia
Better Science , Better Lives.

Phase 2 single arm study of efficacy and safety of P1101 for polycythemia vera (PV) patients for whom the current standard of treatment is difficult to apply.

Development phase:	Phase 2
Document status:	Alive
Release date:	June 07, 2021
Version:	V.3.0
IND No.:	<TBD>
Sponsor:	PharmaEssentia Japan

Revision History

Version	Date	Description
Draft Version 1.0	15/May/2020	Initial Document
Draft Version 2.0	15/Aug/2020	Revision
Draft Version 2.0	15/Sep/2020	Release
Final Version 3.0	07/Jun/2021	Release

1 Signature Page

I have carefully read this statistical analysis plan and agree to the described methods and proceedings.

Author: Phoebe Chen — Date: 08 Jun 2021
Phoebe Chen
Project Statistician
Brightech International, LLC

Brightech Management: Tai Xie — Date: 10 Jun, 2021
Tai Xie, Ph.D.
CEO and Chief Statistician
Brightech International, LLC

PharmaEssentia (Sponsor): Hiroaki Kawase — Date: 11 Jun 2021
Hiroaki Kawase
Clinical Operation
PharmaEssentia Japan Corporation

Table of Contents

List of Abbreviations

Abbreviations	Description of Abbreviations
ADA	Anti-Drug Antibody
AE	Adverse events
AESI	Adverse event of special interest
ATC	Anatomical Therapeutic Chemical
AUC	Area under the time-concentration curve
Cmax	Maximum plasma concentration
CHR	Complete Hematologic Response
CTCAE	Common Terminology Criteria for AEs
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
HU	Hydroxyurea
HCT	Hematocrit
HADS	Hospital Anxiety and Depression Scale
ICH	International Conference on Harmonisation
ITT	Intent-to-treat
ID	Identifier
JAK2	Janus kinase-2
MedDRA	Medical Dictionary for Regulatory Activities

P1101	Ropeginterferon-IFNα-2b
PV	Polycythemia vera
PK	Pharmacokinetic
PPP	Per protocol population
PT	Preferred Term
PLT	Platelet
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SOC	System Organ Class
TEAE	Treatment-emergent adverse event
Tmax	Time to maximum plasma concentration
WBC	White blood cell
WHODrug	World Health Organization Drug Dictionary

2 Overview and Investigational Plan

This Statistical Analysis Plan (SAP) provides a comprehensive and detailed description of the strategy and methodology to be used to perform the analysis of the data generated in the ropeginterferon-IFNα-2b (P1101) polycythemia vera (PV) Japan clinical trial entitled "Phase 2 single arm study of efficacy and safety of P1101 for polycythemia vera (PV) patients for whom the current standard of treatment is difficult to apply." The purpose of the SAP is to ensure appropriate analyses of the study data by using pre-specified statistical approaches prior to the database lock.

This SAP follows the principles of the International Conference on Harmonization (ICH) guidelines E3, E6, and E9 (including E9R1 [1, 2]). This SAP is based on the version 3.0 of the P1101 Japan PV clinical study protocol, dated November 29th, 2019.

2.1 Study Design and Randomization

This is a Phase 2 single arm study to investigate efficacy and safety of P1101 for adult Japanese patients with PV. Eligible patients will be treated with P1101, starting at 100 μg (or 50 μg in patients under another cytoreductive therapy). The dose should be gradually increased by 50 μg every two weeks (in parallel, other cytoreductive therapy should be decreased gradually, as appropriate) until stabilization of the hematological parameters is achieved (hematocrit <45%, platelets <400 x 10^9/L and leukocytes <10 x 10^9/L). The maximum recommended single dose is 500 μg injected every two weeks.

At week 36 (month 9) and week 52 (month 12), the primary study endpoint, phlebotomy-free CHR, will be analyzed. After completion of the 52-week study duration, provision and administration of P1101, collection of the long-term follow up information (blood parameters, molecular and cytogenetic data, safety parameters and as also the optional bone marrow data) will be continued until the drug becomes commercially available for all study subjects.

2.1.1 Randomization and Blinding

This is a single-arm study. No randomization and blinding are needed.

2.2 Study Objectives

2.2.1 Primary Objective

The primary objective is to demonstrate safety and efficacy of P1101 in terms of complete hematologic response (CHR) without phlebotomy with Japanese PV patients for whom the current standard therapy is difficult to apply. The primary

endpoint “complete phlebotomy-free hematologic response (CHR)” is defined as a condition that meets the following hematological criteria:

(1) hematocrit (<45%) without phlebotomy during past 3 months,

(2) normalization of WBC (≤10 x 10^9/L),

(3) platelets (≤400 x 10^9/L) and require no phlebotomy at month 9 and month 12.

2.2.2 Secondary Objectives

The secondary objectives are:

- to demonstrate safety and efficacy of P1101 in terms of changes in HCT, WBC, PLT, spleen size, molecular and cytogenetic response from baseline, time to requiring no phlebotomy, numbers of phlebotomy required, time to response, duration of response maintenance, ratio of CHR rates of subjects with/without prior treatment with HU and Proportion of subjects without thrombotic events.
- to assess the pharmacokinetics (PK) of P1101 and population PK (PPK) analysis
- to assess the relationship between exposure and key efficacy and safety endpoint using exposure-response (E-R) analysis
- to assess the immunogenicity of P1101

2.3 Study Endpoints

2.3.1 Primary Endpoint

2.3.1.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the proportion of subjects who achieved phlebotomy-free complete hematologic response (CHR) at both month 9 and month 12 without phlebotomies during the previous 3 months.

A patient who meets all of the following criteria at assessment visits (month 9 or month 12) will be categorized as a responder:

- Hematocrit <45% phlebotomy-free (absence of phlebotomy during the previous 3 months)
- Platelet count $\leq 400 \times 10^9$/L,
- WBC count $\leq 10 \times 10^9$/L

All three lab parameters are measured by central lab.

If one or more criterion is not met at the assessment visits, the patient will be classified as "non-responder". Patients withdrawn prior to the assessment visit, for any reason, will be classified as “non-responder” for that specific visit.

2.3.2 Secondary Endpoint (s)

The secondary endpoints include:

- Changes in HCT, WBC, PLT and Spleen Size from baseline
- Time during requiring no phlebotomy
- Numbers of phlebotomy
- Time to first peripheral blood count response
- Duration of peripheral blood count response maintenance: Only subjects with response will be included.
- Change from baseline of *JAK2* allelic burden over time
- Spleen size measurements by sonography will be summarized by scheduled time points, as appropriate.
- Difference in CHR rate between subject with or without of treatment with HU
- Proportion of subjects without thrombotic or hemorrhagic events

2.3.3 Safety Endpoints

- Adverse events will be evaluated as follows:

Incidence, causality, and severity of AEs, according to Common Terminology Criteria for Adverse Events (CTCAE; version 5.0)

Incidence, causality, and severity of AEs of special interest (AESIs; e.g., cardiovascular events, hemorrhagic, thrombotic events, psychiatric events)

Events leading to dose reduction or permanent treatment discontinuation

- Other Safety Endpoints

Vital Signs, Clinical Laboratory Tests, Physical Examinations, Electrocardiograms (ECGs), heart echocardiogram, Lung X-rays, Eastern Cooperative Oncology Group (ECOG) performance status, Heart Echo and Ocular Examinations.

2.3.4 Other Endpoints

2.3.4.1 PK Endpoints

- PK parameters of P1101 in plasma samples, including (but not limited to) time to maximum plasma concentration (Tmax), maximum plasma concentration (Cmax), and area under the time-concentration curve (AUC0-t).

  PPK analysis will be conducted based on a separate SAP for the PK data of this study.

2.3.4.2 Immunogenicity Endpoints

- Proportion of subjects who are ADA positive and negative will be determined

2.3.4.3 Exploratory Endpoint

- Changes from baseline in the bone marrow, based on analysis of megakaryopoiesis (normal or increased; patterns of distribution, cellular shape, nuclear shape), age-related cellularity, granulopoiesis and erythropoiesis (decreased/normal/increased, presence of maturation defects), G/E ratio, fibers; % of CD34+ cells; other bone marrow abnormalities.

  *Bone marrow histological remission defined as the disappearance of hypercellularity, trilineage growth (panmyelosis) and absence of >grade 1 reticulin fibrosis (optional)

- Molecular response

2.3.5 Multiplicity Adjustment

No multiplicity adjustment will be considered for this single arm study.

2.4 Determination of Sample Size

Sample size is calculated, based on the following assumptions;

- The target population of the Japanese local study: Japanese PV patients for whom the current standard of treatment is difficult to apply.
- Study design: P1101 single arm no controlled study.
- Primary study objective: To evaluate the safety and efficacy of P1101 for Japanese PV patients for whom the current standard of treatment is difficult to apply.
- Primary efficacy variable: Complete Hematological Response (CHR)
- Study period and evaluation visits: 3 months, 6 months, 9 months and 12 months. The safety follow-up will be continued to 12 months and the major efficacy data will be obtained at 12 months.
- The primary analysis of the primary endpoint: The lower limit of the 95% confidence interval of CHR rate is greater than 11%, and also the slope of the linear regression model of the CHR rate time over the course of the study is greater than 0 (positive trend):
- The rationale: the natural healing rate (placebo response) of CHR rate of PV patients is 0%, thus if the lower limit of 95% confidence interval of the CHR rate exceed 11%, the efficacy of P1101 can be confirmed. The rationale to set the CHR rate limit as 11% is, (1) 11% is the minimum CHR rate value of P1101 obtained in the PROUD-PV study, (2) and the trend of the CHR rate is positive, thus 11% can be considered as a limit of the acceptable clinical risks.
- Assuming the above conditions, the power of the study at 9 months will be about 85% with 25 cases (CHR rate of the PROUD-PV study at 3, 6, 9, and 12 months are 11%, 27%, 36% and 43%, respectively). Considering the uncertainty of maximum 5 patients who are different background from PROUD-PV, a total of 30 cases will be reasonable to evaluate both efficacy and safety.

3 General data handling and definitions

3.1 Groups Column

The Group “P1101” will be used in table outputs.

3.2 Study Visits/ Time-Points

Number of week and number of study visit will be included in label of study visits. The labels will be as follows: Screening, Baseline, Week 2 (V2), Week 4 (V3), Week 6 (V4), …, Week 52 (EOT), Premature discontinuation (for End-of-Study visit of premature discontinued patients), FU (for Safety Follow-Up visit). Unscheduled visit will be labels as 'UNS' and presented in chronological order (UNSv.c), v is visit no. and c is the times of unscheduled visit between Visit v to Visit v+1. Unscheduled visit is visit recorded as '※' in EDC.

Assessment visits will be labeled as follows: Week 12 (V7), Week 24 (V13), Week 36 (V19), Week 52 (EOT). Time-window is defined as +/- 3 days of a schedule visit date. Deviation from the time-window will usually not be adjusted, unless specified otherwise.

3.3 Flags of Values

All the required flags will include but not be limited to, the flags of abnormal values (of vital signs, laboratory assessment, physical examination, ocular examination, lung X-ray examination, heart echo), the flag of target values achieved which defined as the Values of Hematocrit <45%, Platelet count <= 400 x 10^9/L, and WBC count <=10 x 10^9/L, and the flags of central assessments abnormal values. The details will be described in Listing Shell.

3.4 Baseline Values

Baseline value is defined as value of the last assessment prior to the first study drug administration.

Unscheduled and repeated assessments related to Visit 1 will be reviewed during Data Review Meeting (DRM) and questionable baseline assessments, if any, will be confirmed.

3.5 Coded Terms and Dictionaries Used

Medical terms and medications will be coded by MedDRA and WHODrug dictionary as follows:

Item Group	Dictionary (version)	Level presented in data listings*	Level used for creation of frequency tables*
Adverse event term	MedDRA (22.0 or more recent)	Reported Term, PT, SOC	PT
Previous/ concomitant medication name	WHO Drug (March 2019 or more recent)	ATC code, Preferred name	
Medical history	MedDRA (22.0 or more recent)	Reported Term, PT, SOC	PT

*PT = Preferred Term, LLT = Lowest Level Term, HLT = High Level Term, HLGT = High Level Group Term, SOC = primary System Organ Class

Level 1 = Anatomical main group, Level 2 = Therapeutic subgroup, Level 3 = Pharmacological subgroup, Level 4 = Chemical subgroup, Level 5 = Chemical substance

Current version of MedDRA and WHODrug dictionaries will be stated in the tables and listings (in title or footnote).

3.6 Study Periods

Screening period lasts from signature of informed consent to the Week 0 (Visit 1).

Treatment period lasts from Visit 1 to the "End of Treatment" visit or premature discontinuation.

End-of-Treatment (EOT) is the last visit in treatment period.

Post-Treatment Safety follow-up is visit lasting from "End of Treatment" visit (V27) or time of study discontinuation. It is also the End of Study Visit.

3.7 Study Day

Relative study day will be calculated as
Sday=date of interest – first dosing date + 1, if date of interest >= the first dosing date.;
Sday=date of interest – first dosing date, if date of interest < the first dosing date.

3.8 Adverse events

Adverse event (AE) is defined as any event recorded on "Adverse Event" page of e-CRF. Events occurred before date when the informed consent has been signed have to be recorded in "Medical History" page of e- CRF.

Serious adverse event (SAE) is any AE that is reported as serious, as indicated on the "Adverse Event" page of e-CRF.

Treatment-emergent adverse event (TEAE) / treatment emergent SAE (TESAE) is any AE / SAE occurred since date of the first study drug product administration (including that time). If onset date of AE is not complete and beginning of AE after the first drug administration cannot be excluded the AE is considered as treatment emergent. An AE, having been absent pretreatment, or worsens relative to the pretreatment state is also considered as treatment emergent. All questionable cases have to be clarified and approved before analysis.

AE leading to permanent treatment discontinuation, AE leading to dose increasing/ decreasing/ interruption and AE of special interest (AESI) are AEs as indicated on the "Adverse Event" page of e-CRF.

3.9 Previous/Concomitant Medication

Previous/concomitant medications are recorded on the 'Previous/Concomitant medications' page of the e-CRF.

Previous medication is defined as any medication that is taken only prior to the first study drug administration. This includes any medication with end date before the first study drug administration.

Concomitant medication is defined as any medication that is taken after the first study drug administration (including medication ongoing at treatment start).

Previous and concomitant medication as per this definition will be listed and tabulated separately. Concomitant medication started after date of last study drug administration will be flagged in the listings and will be tabulated separately.

3.10 Derivation and Algorithms

3.10.1 Duration of AEs

Variable used for calculation:

Variable Name \ Description	
AESTDAT	Start date
AEENDAT	End date

Duration of AEs (number of days) will be calculated from date of AE onset and AE end using

```
AE_DiffDays=AEENDAT – AESTDAT;
```

3.10.2 Duration of treatment

Variable used for calculation:

Variable Name \ Description
EXDAT \ Date and Time of P1101 Administration

Duration of treatment (in days, weeks, months) will be calculated as follows:

```
TRT_DurDays =
max(EXDAT of last administration)- min(EXDAT of first administration + 14);

TRT_DurWks = TRT_DurDays/7;
TRT_DurWks = TRT_DurDays/30.5;

LABEL
TRT_DurDays = 'Treatment duration [Days]'
TRT_DurWks = 'Treatment duration [Weeks]'
TRT_DurMts = 'Treatment duration [Months]'
;
```

3.10.3 Exposure of P1101 Administration

Variable used for calculation:

Variable Name \ Description
EXSTDTC = min (EXDAT, Date and Time of P1101 Administration)
EXENDTC = max (EXDAT, Date and Time of P1101 Administration)
Treatment Exposure [days] = EXENDTC – EXSTDTC + 14,
Maximum Dose Level [ug]: sorting by unique subject id and use the proc sql to choose the maximum dose level in each subject,
Cumulated Dose [ug]: sum of all used dose by subject,
Mean Dose Level [ug/4 weeks]: 28 x Cumulated Dose [ug] / Treatment Exposure [days]

3.10.4 Interpretation of HADS scores

Total score of depression and anxiety is interpreted as follows:

- Score 0-7: Normal
- Score 8-10: Mild
- Score 11-14: Moderate
- Score 15-21: Severe

4 Statistical and Analytical Procedure

4.1 Analysis Populations

4.1.1 Enrolled Population

The enrolled population will include subjects who have signed an informed consent form.

4.1.2 Intent-to-Treat Population (ITT)

The ITT population will include all treated subjects.

4.1.3 Per Protocol Population (PPP)

The PPP will consist of subjects included in the ITT who are exposed to treatment and have all measurements needed for assessment of the primary endpoint for at least 1 post-baseline visit, and do not have a major protocol deviation.

4.1.4 Safety Population

The Safety Population will include all subjects who have received at least one dose of study drug. For this single arm study, ITT and safety populations are identical.

4.2 Protocol Deviation

The following are considered potential protocol deviations:

- No diagnosis of PV.
- Prior therapy does not match with protocol requirements in terms of types of previous therapy regimens.
- No baseline spleen size measurement in the protocol-specified time windows.
- Subjects who take <80% of the prescribed dose over the total duration of study drug dosing as calculated by the Sponsor based on the drug accountability documented by the site staff and monitored by the Sponsor/designee.

Other observations made during the study will also be considered. All potential protocol deviations will be reviewed prior to database lock. For every potential protocol deviation, a general decision will be made whether to regard it as major or minor, and approved by the Sponsor before database lock. The PPP will be used for efficacy sensitivity analysis.

4.3 Statistical Analysis

This section describes the data analysis in detail. The statistical methods are planned in accordance with the study protocol (Section 17) and in accordance with ICH E9 Statistical Principles for Clinical Trials.

4.3.1 General Principles

Statistical package SAS (version 9.4 or higher) will be used for analysis and for generation of tables, figures and listings (TFL).

In general, data will be presented under group "P1101" as well as overall and by time-point. All data included in the study database will be listed.

4.3.1.1 Descriptive statistics

For categorical data (e.g. response to treatment: responder) standard set of summary statistics will be counts and percentages. Missing, unknown, not available, not applicable or not done categories will be included; therefore, the percentages will be calculated of total number of patients in the set. This approach was selected because it is close to intent-to-treat principle.

For continuous data (e.g. level of lymphocytes) the following descriptive statistics will be presented: number of available observations, mean, standard deviation (SD), median, interquartile range (IQR: Q1-Q3), minimum, and maximum.

Confidence intervals will be present if appropriate.

For time-to-event data (e.g. Time to first peripheral blood count response) KM method will be used for estimating the distribution, median with 95% CI will be presented.

4.3.1.2 Patients' data listings

The following columns will be used for identification of the patients in each listing (whenever applicable):

- Patient ID and Week (Visit).
- For listings of abnormal values (e.g. abnormal values of laboratory assessments, vital signs) and for listing of adverse events the last administered dose (dose administered/prescribed at previous visit) before blood sampling, examination or start of AE will be presented in the listing.

Data listings will be sorted by 'Patient ID' and in chronological order of visits (including unscheduled visits).

4.3.1.3 Rounding procedures

The following rounding procedures will be used:

- Min, Max: same number of decimal places as original value
- Mean, Median and its confidence intervals, Q1, Q3: one more decimal place than original value
- SD: two more decimal places than original value
- Percentages:
  - Safety variables: one decimal place
  - Efficacy variables: two decimal places
  - In case that percentages are very close to 0 or 100: more decimal places can be used.

4.3.1.4 Unscheduled/ repeated assessments

Unscheduled or repeated assessments performed due to safety reasons will only be listed with an exception of assessments which are the last assessments before the first study drug product administration - those will be used as baselines.

4.3.1.5 Missing data and handling of prematurely discontinued patients

Missing data includes missing assessments as well as data missing due to premature termination. No imputation methods are planned to be applied to study data.

4.3.1.6 Methods for handling of incomplete dates/ times

Dates of Adverse Events

For calculation of adverse events duration and for summaries of adverse events by 3 months periods incomplete date and times will be imputed as follows:

Dates of Onset: If AE is recorded at the same month and year when the study drug product was firstly administered to patient then onset date of AEs will be imputed by date of the first administration; otherwise onset date of AEs will be imputed by date of the first day in corresponding month or year.

Dates of End: End dates of AE will be imputed by date of the last day in corresponding month or year.

Dates of Concomitant Medications, adverse events and changes in dosing have to be cleaned before DB lock for final analysis. Hence the questionable cases will be solved individually if needed. The solution will be documented in data review meeting minutes.

4.3.1.7 Validation of statistical programming

The analyses of the primary endpoint, secondary efficacy endpoints and frequency tables of adverse events will be essential to perform double programming. The others also double programmed as applicable.

Logs of all programs used for analysis and data preparation will be checked for errors and unexpected warnings.

Any undocumented updating of study data in statistical programs instead of change in Clinical Database (DB) (or source data) is not allowed. Specifically, this refers to the cases where patients or the data are added/ changed using a statistical program rather than updating the DB. This kind of hard coding is usually proposed to correct deficiencies (missing values, wrong values, and wrong units) in the DB when these errors are detected after DB lock. No hard coding will be done.

This policy ensures integrity of clinical data, since no changes are made to the study data without appropriate documentation from the investigator sites and appropriate audit trails within the clinical trial DB.

4.3.2 Subject Disposition

The number and percentage of subjects in each disposition category (e.g., Enrolled Population, ITT Population, PPP, Safety Population, and Completed Study, or Early Withdraw: with a breakdown of the reasons for early discontinuation) will be summarized by the subjects are treated to. Analysis will be performed using Enrolled Population.

4.3.3 Demographics and Baseline Characteristics

Demographics and Baseline characteristics (e.g., age, sex, weight, height, body mass index, disease history and prior therapies) will be summarized under group "P1101". Analysis will be performed using ITT Population.

4.3.4 Efficacy Variables and Analysis

4.3.4.1 Efficacy Variable

The primary efficacy endpoint is the proportion (CHR) of subjects who achieved the phlebotomy-free complete hematologic response at both month 9 and month 12 without phlebotomies during the previous 3 months. Specifically, subjects who meet all the following criteria at months 9 and 12 will be classified as responder:

- Hematocrit <45% phlebotomy-free (absence of phlebotomy during the previous 3 months)
- Platelet count $\leq 400 \times 10^9$/L,
- WBC count $\leq 10 \times 10^9$/L.

Subjects, who fails to meet all the criteria or withdraw early for any reason, will be classified as non-responder.

4.3.4.2 Secondary Variables

- Changes in HCT, WBC, PLT and Spleen Size from baseline

  Absolute values and changes will be computed over study timepoint (Months 3, 6, 9 and 12)

- Time to first phlebotomy

  Time to first phlebotomy is defined as the time interval from first dosing to first time when subject experiences a phlebotomy procedure. Subjects who completed the study and never had a phlebotomy procedure will be censored at the end of the study (Month 12). Subjects who discontinued from the study prior to any phlebotomy procedure will be censored at the time of discontinuation.

- Numbers of phlebotomy

  Number of phlebotomy procedures experienced will be categorized as 0, 1, 2, 3 and >3.

- Disease response over time (DRT)

  DRT will be derived from hematology parameters (Hematocrit <45% without phlebotomy (at least 3 months since last phlebotomy), and platelets <= 400 x 10^9/L, and WBCs <= 10 x 10^9/L). Central laboratory assessments will be used for disease response assessment at Weeks 12, 24, 36, 52. Patients, who withdraw early for any reason, will be classified as treatment failures in visits that they did not complete. Disease response based on local laboratory assessment will be analyzed as a sensitivity analysis if the data is available.

- Time to first peripheral blood count response

  Time to first peripheral blood count response is defined as the time interval from first dosing to first time when achieving disease response (Hematocrit <45% without phlebotomy (at least 3 months since last phlebotomy), and platelets <= 400 x 10^9/L, and WBCs <= 10 x 10^9/L). Subjects who completed the study and never achieved response will be censored at the end of the treatment (Month 12). Subjects who discontinued prior to achieving a disease response will be censored at the time of discontinuation.

- Duration of peripheral blood count response maintenance:
  Duration of response maintenance will be calculated as number of days starting from assessment visit when the response was achieved till date (excluding) of assessment visit when the response was lost. If more response period occurred within one patient, the longest period will be used for analysis. Number of days will be divided by 30.5 and results presented in months. Only subjects with response will have the duration.
- Change from baseline of *JAK2* allelic burden over time
- Spleen size measurements by sonography will be summarized by scheduled time points, as appropriate.
- Difference in CHR rate between subject with or without of prior treatment with HU
- Proportion of subjects without thrombotic or hemorrhagic events

4.3.5 Statistical Analyses

4.3.5.1 Primary endpoint: CHR

The CHR will be presented as number and proportion of responders who achieved CHR at BOTH Months 9 and 12 along with its 95% CI via exact method (Clopper–Pearson CI).

4.3.5.2 Disease response over time

Disease response over time (DRT) will be summarized over Months 3, 6, 9 and 12 using the same display as for CHR. The slope of the linear regression model of DRT

will be presented. These efficacy variables will be further summarized by prior HU treatment status (with or without).

4.3.5.3 Hematocrit, WBC, Platelets, and Spleen Size change over time

Values of Hematocrit <45%, Platelets count <400 x 10^9/L, and WBC count <10 x 10^9/L will be flagged as target values in data listings.

Count (and percentage) of patients who achieved target values (for Hematocrit, with at least 3 months since last phlebotomy) will be displayed in summary tables together with descriptive statistics. Mean profiles of the parameters will be displayed in graphs.

In listings, all local and central assessment will be presented (and clearly distinguished). In figures, only central assessments will be presented if the parameter was assessed locally and centrally at the certain visit.

Absolute values and changes from baseline will be descriptively analyzed by assessment visits.

4.3.5.4 Number of phlebotomies

Number of phlebotomies during treatment phase will be summarized using descriptive statistics and frequency/percentage for categories of 0,1,2,3 and >3.

4.3.5.5 Time to the first disease response and time to first phlebotomy

KM method will be used for estimating the distribution, median with 95% CI will be presented.

4.3.5.6 Duration of response maintenance

Duration of response maintenance will be calculated as number of days starting from assessment visit when the response was achieved till date (excluding) of assessment visit when the response was lost. If more response period occurred within one patient, the longest period will be used for analysis. Number of days will be divided by 30.5 and results presented in months.

4.3.6 Safety Analysis

Safety data, including ECOG performance status, vital signs, ECGs, heart ECHOs, clinical laboratory tests, physical examinations, lung X-rays, AEs, and ocular examinations, will be summarized by scheduled visits, as appropriate. Change from baseline will be included in summary tables for laboratory, ECG, and vital sign parameters. AEs will be coded by MedDRA and then summarized by system organ class and preferred term for each treatment.

4.3.6.1 Adverse Events

Subjects will be assessed for AEs at each scheduled clinic visit while on the study. All AEs reported on or after the date of first dose until 30 days (inclusive) after the last dose of study drug will be considered treatment-emergent adverse events (TEAEs) and will be summarized by treatment group. For each treatment group, TEAE incidence rates will be summarized with frequency and percentage by System Organ Class (SOC) and preferred term (PT), with safety subjects of that treatment group as the denominator, unless otherwise specified. In addition, TEAE incidence rates will also be summarized by maximum severity (CTCAE Grade)) and relationship to study drug. Treatment-related AEs are those judged by the Investigator to be probably, possibly, or definitely related to the study drug. TEAEs with missing severity or relationship to study drug will be classified as severe and treatment-related, respectively. Subjects with multiple occurrences of events will only be counted once at the maximum severity to study drug for each preferred team, SOC, and overall. Deaths that occur within 30 days after the last dose of study drug are defined as on-study deaths.

Summary tables of the following TEAEs will be provided:

Overall summary of TEAEs: the number and percentage of subjects who experienced any TEAE, severe TEAE, serious adverse event (SAE), any related TEAE, related severe TEAE, related SAE, TEAE leading to treatment discontinuation, and AE leading to death

- TEAEs by SOC and PT
- TEAEs by SOC, PT, and maximum severity
- Related TEAEs by SOC and PT

- TEAEs that led to study drug discontinuation by SOC and PT. Study drug discontinuation will be determined from the AE case report form where action taken is drug withdrawn or etc.
- Serious TEAEs by SOC and PT
- General overview of TEAEs by 3 month periods

4.3.6.2 Physical Examination

Results of physical examination by visit will be reported in frequency tables.

4.3.6.3 Vital signs, ECG parameters, Safety Laboratory data Lung X-Ray, Heart ECHO, and Ocular examination

Values out of normal ranges will be flagged in data listings. Laboratory values will be presented in conventional units. No formal statistical tests comparing results of vital signs and safety laboratory data will be performed.

Absolute values and changes from baseline will be descriptively analyzed by study visits. Count (and percentage) of values in the normal ranges, abnormal clinically significant and not clinically significant will be displayed in summary tables together with descriptive statistics.

In tables and listings all local and central assessment will be presented (and clearly distinguished). In figures local and central assessment will be presented separately.

For assessments without numeric result (e.g. result positive/negative, normal/abnormal) counts and percentages by visit will be presented in the tables and shift tables will be used for presentation of change between Screening and EOT visit/ V27.

4.3.6.4 Bone Marrow Assessment

The parameters of bone marrow assessment, which is an optional measurement in the study, will be listed.

4.3.6.5 Hospital Anxiety and Depression Scale (HADS) and ECOG performance status

Hospital Anxiety and Depression Scale (HADS) classifies an anxiety and a depression as follows: Normal/Mild/Moderate/Severe.

ECOG performance status is classified into 6 levels Grade 0 (Fully active) - Grade 5 (Dead). Descriptive analysis will include frequencies of patients with individual grades by visit.

4.3.7 Prior and Concomitant Medication

All Medications will be coded using the World Health Organization Drug Dictionary (WHODrug).

Concomitant medication will be summarized by Anatomical Therapeutic Chemical (ATC) Level 2 and Preferred Term (PT) / ATC Level 4.

If a medication start date is at or after the first date of treatment, then a medication will be summarized as concomitant medication regardless of whether the medication end date is missing or not.

If a medication end date is before the first date of treatment, then the medication will be summarized as a prior medication regardless of whether the medication start date is missing or not.

If the medication that is started prior to the first dose and continued after treatment will be summarized as the prior medication and separately as the concomitant medication.

4.3.8 Medical History

Medical History will be coded by Medical Dictionary for Regulatory Activities (MedDRA).

4.3.9 Pharmacokinetics Analysis

PK of P1101 in subjects with PV: PK analyses will be performed and summarized in a separate report where PK parameters, including AUC, Cmax, Tmax, and half-life will be summarized with mean, median, SD, and 95% confidence intervals. The effects of covariates on the PK of P1101 will be investigated.

4.3.10 Exploratory Analysis

4.3.10.1 Molecular Response

Molecular response will be calculated from JAK2 mutant allele burden values as following definition [Barosi et al, 2009]

- Complete response (CR): reduction of any molecular abnormality to undetectable levels.
- Partial response (PR)*: reduction from baseline value ≥50% in patients with < 50% mutant allele burden at baseline or reduction from baseline value ≥25% in patients with > 50% or equal to 50%** mutant allele burden at baseline
- No response: Any response that does not satisfy partial or complete response criteria

* Applies only to patients with a baseline value of mutant allele burden greater than 10%. , i.e. for patients with mutant allele burden less than 10% the patient can have only complete response or no response.
** It is not determined how should be evaluated patients with baseline value equal to 50% in the original publication. The definition is completed in this SAP

Molecular responder is included CR and PR. Patient with Negative JAK2 result at baseline will not included in the analysis.

Summarize the number and percentage of molecular response per visit.

5 Software Documentation

Statistical analyses performed by Brightech International will be carried out using SAS®, release 9.4 or higher (SAS/STAT 12.1, SAS Institute Inc., Cary, NC, USA) or above on a Microsoft® Windows® Server 2013 or subsequent platform.

6 Coding Systems

Domain	Coding System	Reporting Terms
Medical history	MedDRA version xx.x	SOC = MedDRA SOC Preferred term = MedDRA Preferred term

Prior and concomitant medication	WHO Drug Dictionary (Version xxx)	ATC Code Medication Group = ATC Level 2 and ATC Level 4 term Drug Number
Adverse events	MedDRA version xx.x	SOC = MedDRA SOC Preferred term = MedDRA Preferred term

7 References

1. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (2014). Draft (final) concept paper on choosing appropriate estimands and defining sensitivity analyses in confirmatory clinical trials. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/E9__R1__Final_Concept_Paper_October_23_2014.pdf.
2. ICH E9 (R1) addendum on estimands and sensitivity 5 analysis in clinical trials to the guideline on statistical 6 principles for clinical trials. 30 August 2017 EMA/CHMP/ICH/436221/2017 Committee for Human Medicinal Products.

   http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2017/08/WC500233916.pdf.
3. Gisslinger H, Gotic M, Holowiecki J, Penka M, Thiele J, Kvasnicka HM, Kralovics R, Petrides PE; ANAHYDRET Study Group. Anagrelide compared with hydroxyurea in WHO-classified essential thrombocythemia: the ANAHYDRET Study, a randomized controlled trial. Blood. 2013 Mar 7;121(10):1720-1728.
4. Agresti, Alan (2002). Categorical Data Analysis. Hooken, New Jersey: John Wiley & Sons, Inc. pp. 231–232. ISBN 0-471-36093-7.
5. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, Frangakis C, Hogan JW, Molenberghs G, Murphy SA, Rotnitsky A, Scharfstein D, Neaton JD, Shih WJ, Siegel JP, Stern H. The prevention and treatment of missing data in clinical trials. New England Journal of Medicine 2012; 367(14): 1355-1360.
6. Rubin DB. Inference and missing data. Biometrika, Volume 63, Issue 3, 1 December 1976, Pages 581–592, https://doi.org/10.1093/biomet/63.3.581.
7. Ropeginterferon alfa-2b versus standard therapy for polycythaemia vera (PROUD-PV and CONTINUATION-PV): a randomised, non-inferiority, phase 3 trial and its extension study. Lancet Haematol. VOLUME 7, ISSUE 3, E196-E208, MARCH 01, 2020, https://doi.org/10.1016/S2352-3026(19)30236-4.