# A Phase 2, Open-label, Multicenter Study to Evaluate the Efficacy and Safety of BN101/belumosudil in Subjects with Chronic Graft Versus Host Disease (cGVHD) After at Least 1 Prior Line of Systemic Therapy

BN101-201

Statistical Analysis Plan

Sponsor：BK Pharmaceuticals Limited

Agent：BioNova Pharmaceuticals Limited

Version: 1.0

Date：17-Jun-2022

Document History

| Version | Date | Author | Description |
|---|---|---|---|
| 1.0 | 17-Jun-2022 | Linlin Dai | Original Document |
| | | | |
| | | | |
| | | | |

**TABLE OF CONTENTS**

## 1. Abbreviations

| | |
|---|---|
| ADI | Actual dose intensity |
| ADL | Activity of daily life |
| AE | Adverse event |
| aGVHD | Acute graft versus host disease |
| ALB | Albumin |
| ALP | Alkaline phosphatase |
| ALT | Alanine aminotransferase |
| AST | Aspartate aminotransferase |
| BMI | Body mass index |
| BUN | Blood urea nitrogen |
| Ca | Calcium |
| cGVHD | Chronic graft versus host disease |
| CI | Confidence interval |
| Cl | Chloride |
| CNI | Calcineurin inhibitor |
| CR | Complete response |
| CRF | Case report form |
| DoR | Duration of response |
| ECOG | Eastern Cooperative Oncology Group |
| ECP | Extracorporeal photopheresis |
| EOT | End of treatment |
| FFS | Failure free survival |
| GGT | Gamma-glutamyl transferase |
| GLU | Glucose |
| GVHD | Graft versus host disease |
| Hb | Hemoglobin |
| K | Potassium |
| LR | Lack of response |
| LR-M | Lack of response - mixed |
| LR-P | Lack of response - progression |
| LR-U | Lack of response - unchanged |
| MedDRA | Medical dictionary for regulatory activities |
| mITT | Modified intent-to-treatment |
| Na | Sodium |
| ORR | Overall response rate |
| OS | Overall survival |
| PDI | Planned dose intensity |
| PK | Pharmacokinetics |
| PLT | Platelet |
| PPS | Per - protocol set |
| PR | Partial response |
| PROMIS | Patient reported outcomes measurement information |

| | |
|---|---|
| | system |
| PS | Performance status |
| PT | Preferred term |
| PTE | Pre-treatment event |
| QD | Once daily |
| QOD | Once every other day |
| RBC | Red blood cell |
| RDI | Relative dose intensity |
| SAE | Serious adverse event |
| SMQ | Standard MedDRA query |
| Cr | Creatinine |
| SOC | System organ class |
| SS | Safety analysis set |
| TBIL | Total bilirubin |
| TEAE | Treatment emergent adverse event |
| TP | Total protein |
| TTNT | Time to new treatment |
| TTR | Time to response |
| UA | Uric acid |
| WBC | White blood cell |

## 2. Introduction

The objective of the BN101-201 study is to evaluate the efficacy and safety of BN101 in the patients with cGVHD who have previously been treated with one to five prior lines of systemic therapy.

This Statistical Analysis Plan (SAP) describes data-handling and statistical procedures to be used for the analysis and reporting of efficacy and safety data collected under Study BN101-201 and presented in the clinical study report (CSR). This SAP has been developed and finalized prior to locking the clinical database. Pharmacokinetic (PK) analyses will be described in separate documents.

## 3. Study Objective

Primary objective：

- To evaluate the efficacy and safety of BN101/belumosudil in subjects with cGVHD who had previously been treated with at least 1 prior line of systemic therapy.

Secondary objective：

- To evaluate the PK profile of BN101/belumosudil in subjects with cGVHD who had previously been treated with at least 1 prior line of systemic therapy;
- To evaluate other efficacy endpoints of BN101/belumosudil in subjects with cGVHD who had previously been treated with at least 1 prior line of systemic therapy.

Exploratory objective：

- To evaluate changes in the Patient-Reported Outcomes Measurement Information System (PROMIS) Global Health sub-scores of physical and mental functioning

## 4. Study Design

BN101-201 is a phase 2, multicenter, open label, single-arm study in patients with cGVHD who have previously been treated with one to five prior lines of systemic therapy. It intends to enroll 30 Chinese cGVHD patients and to evaluate the efficacy, safety, and PK profile of BN101/belumosudil in Chinese cGVHD patients.

The study enrollment period is estimated to be 6 months. 6 months after the last patient enrollment, the database will be locked and the primary analysis will be conducted. After primary analysis, all patients will be followed up for additional 6 months and a follow up analysis will be conducted 12 months after the last patient enrollment.

Eligible patients who have signed informed consent will be enrolled to receive BN101/belumosudil (every 28-day cycle) until cGVHD progression, intolerable toxicity, initiation of a new cGVHD therapy, recurrence of hematological neoplasm, loss to follow-up,

withdrawal of consent, or death, whichever occurrs first. During study, concomitant treatment with standard of care systemic cGVHD therapies such as corticosteroids, calcineurin inhibitors (tacrolimus, cyclosporine), sirolimus, MMF, or extracorporeal photopheresis (ECP), are permitted as long as the patient has been on a stable dose / regimen of these per the eligibility criteria. Initiation of new systemic cGVHD therapy while on study is not permitted. Response criteria will be assessed after cycles 1, 2, 3 and 4 and then after every other cycle thereafter (i.e. on Day 1 of Cycle 2, 3, 4, 5, 7, 9, 11 etc). For patients assessed as "Lack of Response - Mixed" and "Lack of Response - Progression" [Evaluated by the investigator based on the National Institutes of Health (NIH) Consensus Development Project on Criteria for Clinical Trials in cGVHD(2014), hereinafter referred to as the NIH Consensus Criteria (2014) ], if the investigator believed that these patients could continue to benefit from treatment with BN101/belumosudil, the treatment may be continued after a written application was submitted, and approval from the Sponsor's Medical Monitor as well as documentation of the patient's willingness to continue treatment was obtained. For patients who did not achieve any response after 12 cycles of BN101/belumosudil treatment, they should discontinue the study medication and withdraw from the study if in the investigator's judgment there is no evidence of clinical benefit (e.g., improvements in organ score, improvements in Lee symptom scores, reductions of corticosteroid/tacrolimus doses). For patients whose cGVHD has not progressed at the time of discontinuation of the study treatment and who come off study for reasons other than AEs should be tapered off BN101/belumosudil by reducing the dose every other week (200mg QD → 200mg QOD → discontinue).

# 5. Study Endpoints

## 5.1. Efficacy Endpoints

### 5.1.1. Primary Efficacy Endpoints

The primary efficacy endpoint is the overall response rate (ORR) as assessed by the investigators based upon the 2014 NIH Consensus Development Project on Clinical Trials in cGVHD.

Response is assessed using individual scores from ten systems (Skin, Eyes, Mouth, Esophagus, Upper GI, Lower GI, Liver, Lungs, Joints and Fascia and Global Severity Rating (GSR)). Response is assessed with respect to the baseline (Cycle 1, Day 1 (C1D1)) cGVHD assessment. The overall response at each assessment time point will be categorized as Complete Response (CR), Partial Response (PR), or Lack of Response (LR), where LR includes the response status of unchanged, mixed, or progression as defined in Table 1.

The ORR is defined as the proportion of patients with a best response meeting the overall response criteria assessment of CR or PR at any post-baseline response assessment.

If a treated patient is lost to follow up without response assessment, this patient will be counted as a non-responder. Any response on and after new systemic cGVHD treatment will be censored as non-response.

Table 1: cGVHD Response Definitions

| Response | Definition |
|---|---|
| Complete Response (CR) | Resolution of all manifestations of cGVHD in each organ or site. |
| Partial Response (PR) | Improvement in at least one organ or site without progression in any other organ or site. |
| Lack of Response (LR) | |
| Mixed (LR-M) | Complete or partial response in at least one organ accompanied by progression in another organ* |
| Unchanged (LR-U) | Outcomes that do not meet the criteria for complete response, partial response, progression or mixed response |
| Progression (LR-P) | Progression in at least one organ or site without a response in any other organ or site |
| * Considered progression for the purposes of analysis | |

### 5.1.2. Secondary Efficacy Endpoints

Secondary efficacy endpoints include Duration of Response (DoR), Lee Symptom Scale Score (LSS), Response by Organ System, Time to Response, Time to Next Therapy, Corticosteroid Dose, Calcineurin Inhibitor (CNI) Dose, Failure Free Survival (FFS), Overall Survival (OS), Change in cGVHD severity using the Clinician-Reported Global cGVHD Activity Assessment, and Change in symptom activity using the cGVHD Activity Assessment Patient Self-Report.

**DoR**

Four definitions of DoR are as following:

The primary definition of DoR is the time from first documentation of response to the time of first documentation of deterioration from best response (e.g., CR to PR, or PR to LR).

The secondary definition of DoR is the time from first documentation of response to the time of first documentation of lack of response.

The tertiary definition of DoR is the time from first documentation of response to the time of initiation of new systemic cGVHD therapy.

The quaternary definition of DoR is the time from first documentation of response to the time of first documentation of lack of response (as the secondary definition) but with durations summed for multiple response / lack of response episodes.

The DoR will be reported only for responders. And the censoring rules in Table 2 will be applied.

**Lee Symptom Scale Score (LSS)**

Lee cGVHD Symptom Scale will be assessed on Day 1 of Cycles 2-5, then on Day 1 of every other Cycle thereafter and EOT. The questionnaire consists 30 items of 7 domains: skin, eyes and mouth, breathing, eating and digestion, muscles and joints, energy, and emotional. Each question is scored 0, 1, 2, 3 or 4.

A domain score will be calculated for each domain by taking the mean of all items completed if more than 50% were answered and normalizing to a 0 to 100 scale. A summary score will be calculated as average of all domain scores if more than 50% of them are nonmissing. A higher score indicated more bothersome symptoms. 7 points difference on the summary score of cGVHD symptom scale was found to be clinically meaningful.

*Raw Score* = （$I_1$ + ・・・ + $I_m$ ）/ m， m is the number of items have nonmissing value in a domain.

*Domain Score* = *Raw Score* / 4 *100。

*Summary Score* = （*Domain Score*$_1$ + ・・・ + *Domain Score*$_n$）/ *n*， n is the number of domains have nonmissing domain scores.

**Response by Organ System**

Response assessment by physician for the nie individual organs (skin, eyes, mouth, esophagus, upper GI, lower GI, liber, lungs and joints and fascia) plus global severity rating will be collected in eCRF.

**Time to Response (TTR)**

TTR is defined as the time from first treatment to the time of first documentation of response.

**Time to Next Therapy (TTNT)**

TTNT is defined as the time from first treatment to the time of new systemic cGVHD treatment, censored by last response assessment or long term follow up assessment, which ever is the latest and available.

**Corticosteroid Dose**

Corticosteroid doses will be presented as mg/kg/day prednisone equivalent dose. If a patient is not using prednisone as the systemic corticosteroid, then the prednisone dose equivalent will be determined according to following conversion ratios:

1mg prednisone is equivalent to:

- 4mg Hydrocortisone
- 0.8mg Methylprednisolone
- 0.15mg Dexamethasone

- 1mg Prednisolone
- 0.8mg Triamcinolone

Transient increase in corticosteroid dosing (not exceeding 1mg/kg/day prednisone equivalent) is permitted for the treatment of cGVHD flare, but dose must be reduced back to the pre-enrollment dose within 6 weeks. If the dose remains elevated for more than 6 weeks, this will be considered a BN101/belumosudil treatment failure. More than 2 episodes of cGVHD flare requiring increased corticosteroid in the first month of BN101/belumosudil treatment will also be considered a BN101/belumosudil treatment failure.

**Calcineurin Inhibitor (CNI) Dose**

CNI includes systemic tacrolimus and cyclosporine.

**Failure Free Survival (FFS)**

FFS is defined as the absence of new cGVHD systemic therapy, non-relapse mortality and recurrent malignancy (e.g. underlying disease) and is censored by last response assessment or long term follow up assessment, whichever is the latest and available.

**Overall Survival (OS)**

OS is defined as time from first dose of BN101/belumosudil to the date of death due to any cause.

**Change in cGVHD severity using the Clinician-Reported Global cGVHD Activity Assessment**

The Clinician-reported global cGVHD Activity Assessment is a 0-10 point numeric rating scale with a score 0 indicating “cGVHD symptoms not at all severe” and a score of 10 being “most severe cGVHD symptoms possible”. The activities are assessed on Day 1 of each cycle from Cycle 1 Day 1 to EOT.

**Change in symptom activity using the cGVHD Activity Assessment Patient Self-Report**

The symptom activity item is a 0-10-point numeric rating scale with a score of 0 indicating “cGVHD symptoms not at all severe” and a score of 10 being “most severe cGVHD symptoms possible”. The status reported by patients are categorized as none, mild, moderate, and severe. The comparison of cGVHD symptoms to a month ago will also be reported by patients, rating from -3 (very much worse) to +3 (very much better).

### 5.1.3. Exploratory Endpoints

Patient-Reported Outcomes Measurement Information System (PROMIS), a state-of-the-science PRO measurement system, was developed using mixed qualitative and quantitative methods and uses item response theory-calibrated item banks for numerous patient-reported symptoms and functional domains. The PROMIS Global Health 10 measure comprises ten items with two summary scores for physical and mental functioning with higher scores indicating better

functioning. Those two summary scores will be calculated using the Assessment Center Scoring Service (https://www.assessmentcenter.net/ac_scoringservice) from raw data.

## 5.2. Safety Endpoints

### 5.2.1. Adverse Events

**Adverse Event**

Adverse event (AE) is defined as any untoward medical occurrence in a patient administered the study medication and which does not necessarily have to have a causal relationship with the treatment.

**Treatment Emergent Adverse Event**

Treatment emergent adverse events (TEAE) are any AEs occurring or worsening in severity after the first administration of study medication.

**Significant Adverse Event**

Significant adverse events are any AEs leading to additional pertinent intervention (such as dose interruption, dose reduction or dose discontinuation).

**Causality of Adverse Event**

The causality with BN101/belumosudil will be classified into following 5 classes: definitely related, possibly related, unlikely related, not related, and not determined. Definitely related, possibly related, and not determined will be considered related in the analysis of related AEs.

**Serious Adverse Event（SAE）**

Serious adverse events (SAE) are any AE which happens during any study period (screening period, treatment period, follow up period) with any dose of the study medication and which meet one or more following criteria. Treatment emergent SAE will be summarized in the analysis.

- Result in death
- Life threatening
- Requires inpatient hospitalization or prolongation of existing hospitalization
- Result in persistent or significant disability / incapacity
- Result in a congenital anomaly / birth defect
- Other important medical events

**Death**

All AEs leading to death will be recorded in CRF. Death report will also be collected in CRF.

**Severity of Adverse Events**

Adverse event will be graded using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 5.0：

Grade 1: Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated.

Grade 2: Moderate; minimal, local or noninvasive intervention indicated; limiting age-appropriate instrumental Activities of Daily Living (ADL) (instrumental ADL refer to preparing meals, shopping for groceries or clothes, using the telephone, managing money, etc)

Grade 3: Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self care ADL (self care ADL refer to bathing, dressing and undressing, feeding self, using the toilet, taking medications, and not bedridden)

Grade 4: Life threatening consequences; urgent intervention indicated.

Grade 5: Death related to AE.

### 5.2.2. Vital Signs

Body temperature, pulse, blood pressure (systolic pressure and diastolic pressure) and overall assessment will be collected in CRF.

### 5.2.3. 12 Lead-ECG

12 lead-ECG results (heart rate, RR interval, PR interval, QT interval, QTcB interval, QTcF interval, QRS wave and overall assessment) will be collected in CRF.

### 5.2.4. Laboratory Evaluations

Hematology, chemistry, and urinalysis will be collected in CRF.

### 5.2.5. Eastern Cooperative Oncology Group (ECOG) Performance Status (PS) Score

ECOG PS score will be collected in CRF.

## 5.3. Pharmacokinetic Variables

Pharmacokinetic variables will be presented in a separate PK analysis plan.

## 5.4. Protocol Deviation

All protocol deviations will be identified and classified as major or minor (as defined below) before the clinical database lock.

Major Deviations: Protocol deviation that may impact the accuracy, and/or reliability of the study data or that may impact patient rights, safety or well-being.

Minor Deviations: Protocol deviation that does not impact the accuracy, and/or reliability of the study data or patient rights, safety or well-being.

### 5.5. Treatment Compliance and Exposure

The extent of BN101/belumosudil exposure will be assessed using the duration of BN101/belumosudil treatment.

Duration of treatment (days) = last dose date – first dose date + 1 regardless of unplanned intermittent discontinuations.

The cumulative duration of treatment will be provided, defined as the sum of the duration of treatment for all patients, and will be expressed in patient years.

The relative dose intensity (RDI) will be presented. The RDI is defined as:

RDI (%) = 100% * ADI (mg/day) / PDI (mg/day) 。

Where ADI and PDI are the actual dose intensity and planned dose intensity, respectively:：

ADI (mg/day) = actual cumulative dose (mg) / duration of treatment (days)

PDI (mg/day) = planned cumulative dose (mg) / duration of treatment (days)

The planned cumulative dose is the planned daily dose * duration of treatment, while the actual cumulative dose is the sum of actual doses received over the duration of treatment. The actual total dose will incorporate information of dose reduction and dose interruptions captured in the CRF.

## 6. Statistical Hypothesis

Since this study intend to evaluate the consistency of the efficacy, safety and PK profile of belumosudil between Chinese cGVHD patients and Caucasian cGVHD patients for the purpose of bridging the results of KD025-213 study, no formal hypothesis testing will be performed. All statistical analyses will be descriptive.

## 7. Analysis Population

**Modified Intent-to-treat (mITT)**

mITT: The primary population for efficacy analyses will be a modified Intent to treat (mITT) population defined as all enrolled patients who receive at least one dose of study medication.

The mITT population will be used for tables of demography, baseline characteristics and efficacy.

Responder: The responder population is defined as patients in the mITT population that achieved a partial or complete response at any post-baseline response assessment.

Nonresponder: The nonresponder population is defined as anyone in mITT population that is not a responder.

The responder and nonresponder populations will be used for some subgroup analyses.

**Safety Set (SS)**

The safety population is defined as all enrolled patients who receive at least one dose of study medication and have at least one post-baseline safety assessment. Safety population will be used for all safety analyses.

**Per-protocol Set（PPS）**

The per-protocol population is defined as a subset of mITT, including all enrolled patients who receive at least one dose of study medication without major protocol deviations. PPS will be used for sensitivity analyses of primary efficacy analysis.

## 8. Statistical Methods

### 8.1. General Consideration

In the line listing, subject ID (site ID-patient screening number) will be the only universal identifier for the patient.

General rule for the statistical description: continuous variables will be summarized by descriptive statistics of mean, median, standard deviation, minimum and maximum. Categorical variables will be summarized with counts and corresponding percentages.

All analyses will be conducted using SAS version 9.4 or higher.

**Baseline**

Baseline is defined as the last non-missing measurement or assessment taken on or prior to first administration of the study medication.

**Study Day**

Study day will be calculated from the date of first administration of the study medication.

If the date of event / assessment / sample collection is prior to the first administration of the study medication:

Study day = date of event / assessment / sample collection – date of first administration of the study medication

The day before the date of first dose will be study day -1.

If the date of event / assessment / sample collection is after the first administration of the study medication:

Study day = date of event / assessment / sample collection – date of first administration of the study medication +1

The day of the first administration of the study medication is study day 1.

Unless otherwise specified, unscheduled measurements will not be included in the by-visit summaries, but will contribute to best / worst case value, minimum / maximum changes, or abnormal or significant observations where required (e.g. shift table).

## 8.2. Handling of Data

### 8.2.1. Incomplete or Missing Data

Unless otherwise specified, the missing data will not be imputed. If the missing data is categorical, it should be classified as "missing" or "unknown".

Incomplete or Missing Start Date and End Date of AE, Prior / Concomitant Medication, Prior / Concomitant Non-medication Treatment

Imputation of missing start date and end date of AE, prior/concomitant medication, prior/concomitant non-medication treatment is intended to decide the precedence relationship of the event relative to the first administration of the study medication. Unimputed date will be presented in the listing.

- Incomplete or Missing Onset Date

If onset year and month are known and onset year and month are prior to the year and month of first dose of study treatment, the onset date will be imputed as the 1st day of onset month. If onset year and month are equal to the year and month of first dose of study treatment, then the onset date will be imputed as the date of first dose of study treatment.

If only onset year is known, then the onset date will be imputed as January 1st. If onset year is equal to the year of first dose of study treatment, then the onset date will be imputed as the date of first dose of study treatment.

If onset year, month and date are all missing, then the onset date will be imputed as the date of first dose of study treatment.

Note: when the onset date is imputed, it should be on or prior to the end date of event.

- Incomplete or Missing End Date

If end year and month are known, the end date will be imputed with the last day of end month. If the imputed date is later than the latest assessment date, then the end date will be imputed as the latest assessment date.

If end year is known, the end date will be imputed as the December 31st. If the imputed date is later than the latest assessment date, then the end date will be imputed as the latest assessment date.

If end year, month and date are all missing, it should be determined whether the event is ongoing. If the variable of ongoing or not is yes or missing, then the end date will not be imputed. If the variable of ongoing or not is not, then the end date will be imputed as the latest assessment date.

Incomplete or Missing Event Date in Time to Event Analyses

In time to event analyses, if the year and month of event are known, the event date will be imputed conservatively. If the event is not anticipated, then the event date will be imputed as the 1st day of the month. If the event is neutral, then the date will be imputed as the 15th of the month. If the event is anticipated, then the event date will be imputed as the last day of the month.

### 8.2.2. Derived and Transformed Data

Data will not be derived or transformed except the necessary calculation for end points based on data collected in CRF. Regarding the outliers, queries will be raised to data manager and they should be resolved or confirmed before database lock.

## 8.3. Study Population

### 8.3.1. Patient Disposition

A disposition of all enrolled patients will be summarized. The number and percentage of patients discontinuing from the study and the primary reason for discontinuation will be summarized.

### 8.3.2. Protocol Deviation

Major protocol deviations will be categorized by deviation type (e.g. Entry/eligibility, use of excluded medication). The subject incidence of major protocol deviations will be summarized overall and by deviation category. All major protocol deviations will be presented in a listing as well.

### 8.3.3. Demographic

Patient demographics and baseline characteristics will be summarized for the mITT population. Descriptive statistics will be provided for age, height, weight, and body mass index (BMI). Frequencies and percentages will be tabulated for sex, ethnicity, and Eastern Cooperative Oncology Group (ECOG) performance status (PS) score.

### 8.3.4. Medical History

Medical history will be summarized by primary system organ class (SOC) and preferred term (PT). Medical history will be coded using Medical Dictionary for Regulatory Activities (MedDRA, Version 25.0 or higher) terminology. Frequencies and percentages will be tabulated for medical history by SOC and PT.

### 8.3.5. Transplant History

Transplant history will be summarized for:

- Auto-transplant (Yes/No)
- Auto-transplant number
- Indication for transplant
- Allo-transplant number
- HLA match (matched, unmatched)
- Stem cell type (bone marrow, peripheral blood, bone marrow + peripheral blood, cord blood)
- Donor source (related, unrelated)
- Conditioning regimen (myeloablative, nonmyeloablative)
- Donor age
- Donor gender
- Donor CMV serostatus
- Recipient CMV serostatus
- Prophylaxis for GVHD

### 8.3.6. GVHD History

The baseline characteristics of GVHD will be summarized for:

- Prior aGVHD
- Time from transplant to cGVHD diagnosis
- Time from cGVHD diagnosis to enrollment
- Time from transplant to enrollment
- cGVHD severity: mild / moderate / severe
- Number of organs involved at baseline
- Organ involvement at baseline per NIH criteria

### 8.3.7. Prior Therapies for cGVHD

Prior therapies for cGVHD will be summarized for:

Number of prior lines of systemic cGVHD therapy

Prior systemic cGVHD therapies

Best response to the recent systemic therapy

### 8.3.8. Treatment Compliance and Exposure

Treatment compliance and exposure will be analyzed on mITT population.

Duration of treatment and duration of follow up will be summarized as a continuous variable and also categorically by numbers and percentages for each category (0-3 months, 3-6 months, 6-9 months, >9 months).

Descriptive summary statistics of cumulative exposure (patient year), actual cumulative dose (mg), ADI (mg/day), RDI (%) and RDI as a categorical endpoint (>95%, ≤95%) will be presented.

Number and percentage of patients for each RDI category (<80%, 80-100%, >100%) will be summarized by visit.

Number and percentage of patients with dose reduction and dose interruption will be summarized.

## 8.4. Efficacy Analyses

Analyses for primary endpoint will be based on mITT and PPS. Other efficacy analyses will be based on mITT unless otherwise specified.

### 8.4.1. Analyses for Primary Endpoints

ORR（%）=（*the number of patients with a best response meeting the criteria of CR or PR at any post-baseline response assessment / the total number of patients in the analysis set*）*100%。

ORR will be calculated based on mITT and PPS. Point estimate and 95% confidence interval and 97.5% confidence interval (Clopper-Pearson (exact) method and Normal Approximation method) of ORR will be provided.

### 8.4.2. Analyses for Secondary Endpoints

**DoR**

DoR will be analysed only based on responder population. Kaplan-Meier plots and descriptive statistics of DoR will be provided. The censoring rules in Table 2 will be applied. Landmark analyses with number and percentage of patients with response sustained for ≥12, ≥16, ≥20, ≥24, ≥32, ≥40 and ≥48 weeks will be provided.

Table 2 Censoring Rules for Duration of Response

| DoR | Events | Censoring |
|---|---|---|
| Primary | • Deterioration from best response<br>• Initiation of new systemic therapy for cGVHD<br>• Death<br>If deterioration from best response or initiation of new systemic therapy happens immediately after two or more missed response assessments, the event date should be set as four weeks (one cycle) after last documented response assessment prior this event. | Last documented response assessment |
| Secondary | • Documented LR<br>• Initiation of new systemic therapy for cGVHD<br>• Death<br>If LR or initiation of new systemic therapy happens immediately after two or more missed response assessments, the event date should be set as four weeks (one cycle) after last documented response assessment prior this event. | |
| Tertiary | • Initiation of new systemic therapy for cGVHD<br>• Death | Last response assessment or long term follow up assessment, whichever is the latest and available. |
| Quaternary | • LR<br>• Initiation of new systemic therapy for cGVHD<br>• Death<br>With summation of DoR from multiple episodes. | Same with censoring rule for primary and secondary. |

**Lee Symptom Scale Score (LSS)**

The analyses will be performed based on mITT, responder and nonresponder population. The analyses will include:

- Both scores and change from baseline values (summary score and domain scores) will be summarized as continuous variables by visit.
- Number and percent of patients with a ≥7 point reduction (7-PtR) from baseline.

- Number and percent of patients with a 7-PtR from baseline on 2 consecutive assessments.

- Duration of a 7-PtR (DO7-PtR) (defined as time from documentation of the first ≥7-point resuction to the first documentation of less than 7-point reduction). If there are multiple episodes, the DO7-PtR will be measured as the sum of DO7-PtR from all episodes.

**Response by Organ System**

The best response (CR, PR) for the nine individual organs plus global severity rating will be summarized. Time to response at the organ level will also be evaluated. Descriptive statistics and plots of cumulative number and percentage of responders over time will be provided. Two series of percentages will be presented:

- With total number of patients in mITT with involvement of the given organ at baseline as denominator

- With numbers of patients in the responder population with involvement of the given organ at baseline as denominator

**Time to Response (TTR)**

Descriptive statistics and plots of cumulative number and percent of responders over time will be provided. TTR analyses will only be conducted for the responder population.

**Time to Next Therapy (TTNT)**

TTNT will be analysed by the Kaplan-Meier survival method as well as the landmark analysis at 3, 6, 9, and 12 months.

**Corticosteroid Dose**

Corticosteroid doses will be presented as prednisone equivalent dose. Descriptive statistics for the mITT population, responder / nonresponder populations will be provided for:

- Systemic corticosteroid dose over time

- Change and percent change from baseline (C1D1) to the greatest corticosteroid dose reduction during BN101/belumosudil treatment period

- Number and percentage of patients who reduced systemic corticosteroid dose during BN101/belumosudil treatment period

- Number and percentage of patients who ever discontinued systemic corticosteroid usage for over 28 days during BN101/belumosudil treatment period

The above analyses will be censored on or after new systemic cGVHD treatment.

**Calcineurin Inhibitor (CNI) Dose**

CNI includes systemic tacrolimus and cyclosporine. Descriptive statistics will be provided for:

- Number and percentage of patients who reduced CNI dose during BN101/belumosudil treatment period
- Number and percentage of patients who ever discontinued CNI during BN101/belumosudil treatment period

**Failure Free Survival (FFS)**

Kaplan-Meier plots, descriptive statistics of FFS and landmark FFS at 3, 6, 9, and 12 months will be provided. In addition, analyses for the three components of FFS will also be provided.

**Overall Survival (OS)**

Kaplan-Meier plots, descriptive statistics of OS and landmark OS at 3, 6, 9, and 12 months will be provided.

**Change in cGVHD severity using the Clinician-Reported Global cGVHD Activity Assessment**

Change from baseline in cGVHD severity based on the Clinician-reported global cGVHD Activity Assessment will be summarized as a categorical endpoint at all scheduled assessment visits.

**Change in symptom activity using the cGVHD Activity Assessment Patient Self-Report**

Changes in cGVHD symptoms based on global cGVHD Activity Assessment by the Patient Self-Report will be summarized as a continuous endpoint at all scheduled assessment visits. Both scores as well as the change-from-baseline values will be presented.

The summary of the change-from-baseline of Global Severity Rating on categorical status and the summary of comparison of cGVHD to a month ago will also be presented.

The number and percentage of patients reporting none, mild, moderate, and severe cGVHD will be summarized by visit.

### 8.4.3. Additional Analyses

**Exploratory Analyses – PROMIS Global Health Scores**

The PROMIS Global Health 10 measure comprises ten items with two summary scores for physical and mental functioning. The analyses will be performed on mITT, responder and nonresponder populations. The analyses will include:

- Both raw scores and change-from-baseline values (physical and mental domains) will be summarized as continuous variables by visit
- Number and percentage of patients with a ≥4.7-point reduction from baseline (C1D1)
- Number and percentage of patients with a ≥4.7-point increase from baseline (C1D1)

### 8.5. Safety Analysis

Safety analyses will be conducted on the safety analysis set. Safety analyses will include AEs, SAEs, vital sign measurements, clinical laboratory evaluations (hematology, chemistry) and ECGs.

#### 8.5.1. Adverse Events

AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) dictionary (Version 25.0 or higher). All AEs (including SAEs) will be graded using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 5.0.

Summary of overall treatment emergent adverse events will be tabulated. The subject incidence of the following TEAEs will be presented: any TEAEs, any TEAEs with worst severity, treatment related TEAEs, serious TEAEs, treatment related serious TEAE, ≥grade 3 TEAEs, ≥grade 3 treatment related TEAEs, ≥grade 3 serious TEAEs, ≥grade 3 serious treatment related TEAEs, TEAEs leading to dose interruption, TEAEs leading to dose reduction, TEAEs leading to treatment discontinuation, and TEAEs leading to death.

Following TEAEs will be tabulated by System Organ Class (SOC) and Preferred Term (PT): all TEAEs, serious TEAEs, ≥grade 3 TEAEs, treatment related TEAEs, ≥grade 3 treatment related TEAEs, and serious treatment related TEAEs. Patient will be counted only once within a SOC and PT.

Following TEAEs will be tabulated by PT: all TEAEs, treatment related TEAEs, serious TEAEs, serious treatment related TEAEs, TEAEs leading to dose reduction, treatment related TEAEs leading to dose reduction, TEAEs leading to dose interruption, treatment related TEAEs leading to dose interruption, TEAEs leading to treatment discontinuation, treatment related TEAEs leading to treatment discontinuation, ≥grade 3 TEAEs, ≥grade 3 treatment related TEAEs, CMQ bacterial infection, CMQ viral infection, CMQ fungal infection, CMQ cytopenia, SMQ liver related investigations, signs and symptoms, SMQ drug related hepatic disorders – comprehensive search (narrow), ≥grade 3 CMQ bacterial infection, ≥grade 3 CMQ viral infection, ≥grade 3 CMQ fungal infection, ≥grade 3 CMQ cytopenia, ≥grade 3 SMQ liver related investigations, signs and symptoms, and ≥grade 3 SMQ drug related hepatic disorders – comprehensive search (narrow). Patient will be counted only once within a PT.

Subject listings will be provided for AEs, SAEs, AEs leading to dose interruption, AEs leading to dose reduction, AEs leading to treatment discontinuation, and deaths.

#### 8.5.2. Laboratory Evaluations

The summary statistics of all hematology and biochemistry variables and change from baseline will be presented for each study assessment.

For shift tables, laboratory results will be classified using the CTCAE V5.0. All graded laboratory parameters will be summarized separately for hematology and biochemistry. Corresponding shift tables to compare baseline to the worst post-baseline grade within the treatment period will be provided.

For parameters of hematology and biochemistry, plots of mean / mean change from baseline with the corresponding standard error will be displayed.

Subject listing for hematology, biochemistry and urinalysis will be provided.

### 8.5.3. 12 Lead ECG

Descriptive statistics for ECG parameters at each time point will be presented for the values and change from baseline.

The number and percentage of patients with observed QTcF values that satisfy the following conditions will be presented by study visit. The number and percentage of patients with longest QTcF values that satisfy the following conditions will be summarized.

- ≤ 450 msec
- >450 to ≤ 480 msec
- >480 to ≤ 500 msec
- >500 msec

The number and percentage of patients having change from baseline QTcF values that satisfy the following conditions will be presented by study visit. The number and percentage of patients with longest change from baseline QTcF values that satisfy the following conditions will be summarized.

- ≤ 0 msec
- >0 to ≤ 30 msec
- >30 to ≤ 60 msec
- > 60 msec

Subject listing for ECG results will be provided.

### 8.5.4. Vital Signs

Descriptive statistics for vital signs (weight, BMI, body temperature, respiratory rate, pulse rate, systolic and diastolic blood pressure) values and the change from the baseline will be presented for each scheduled assessment time point.

Subject listing for vital signs will be provided.

### 8.5.5. Physical Examination

Clinically significant physical examination findings will be captured as AEs.

## 8.6. Pharmacokinetic Analyses

PK analyses will be presented in a separate PK analysis plan.

## 8.7. Other Analyses

Subgroup Analyses

Subgroup analyses for the endpoint of ORR will be conducted for the following subgroups：

- Severe cGVHD at screening (Yes / No)
- Number of organs involved at baseline (<4 / ≥4)
- Number of prior lines of therapy (1 / ≥2)
- Duration of cGVHD before enrollment (by $50^{th}$ percentile)
- Lung involvement at baseline (Yes / No)
- Prior Ruxolitinib (Yes / No)
- Prior ibrutinib (Yes / No)

# 9. References

NMPA : Statistical Principles for Clinical Trials. 2016

NMPA : Guidance for Data Management and Statistical Analysis Plan and Report in Clinical Trial. 2016