

Clinical Development


BKM120 (Buparlisib)

Study Number: CBKM120H2201 / NCT01852292

Phase II multicenter randomized, double blind, placebo controlled study assessing the efficacy of buparlisib (BKM120) plus paclitaxel vs. placebo plus paclitaxel in patients with platinum pre-treated recurrent or metastatic head and neck squamous cell carcinoma

RAP Module 3 – Detailed Statistical Methodology

Amendment 2

Author: 

Document type: RAP Documentation

Document status: Final v3.0

Release date: 22 Apr 2016

Number of pages: 55

Document History – Changes compared to previous version of RAP module 3.

Version	Date	Changes
0.1	09-12-2013	Initial version for RAP meeting I
1.0	09-01-2014	Final version
2.0	15-Sep-15	<p><i>Sec 2.1,2.2,2.5</i></p> <p><i>Alignment with MAP</i></p> <ul style="list-style-type: none">- Definition of Baseline, Last contact date, Screened Patients, aligned with current standard- Handling of data analyses in case of withdrawal of consent aligned with current standard- Pharmacokinetics analysis set and Full Pharmacokinetic analysis set added due to addition of PK in protocol amendment 1- Per protocol set definition updated to exclude patients who were randomized but not treated.- PK analyses added based on protocol amendment 1- BOR-unknown reason SD too early clarified as per current standard-Determination of missing adequate Tumor assessments – definition of D1 modified to 31 days.- Included summarizing of Unknown/Missing Tumor assessments by treatment groups. <p><i>Sec 3.5</i></p> <ul style="list-style-type: none">-It is clarified that HPV status will be derived using central lab data and historical HPV status will be listed.-Sec 3.5.7 : Added an additional summary of prior antineoplastic by number of lines in Recurrent/Metastatic setting. Drugs being classified as EGFR inhibitors are mentioned in M8. <p><i>Sec 3.6</i></p> <ul style="list-style-type: none">- Added a dose window for Paclitaxel to take into account the variation of total dose due to BSA changes- Updated cumulative dose calculations for BKM intermitent dosing based on current program standards- Updated dose interruptions / reductions based on current program standards <p><i>Sec 3.8.</i></p> <ul style="list-style-type: none">-As per protocol amendment 2, it is mentioned that primary analysis will be done after approximately 120 PFS events are documented. Final Analysis will be done after approximately 112 Deaths are observed. <p><i>Sec 3.8.1.2</i></p> <ul style="list-style-type: none">-Updated the methodology for median time to censoring (Schemper, M ; Smith, TL , 1996) <p><i>Sec 3.8.1.3</i></p> <ul style="list-style-type: none">-Sensitivity analyses added to repeat primary analysis based on stratification variables derived from eCRF in addition to the primary analysis based on IRT strata.

Version	Date	Changes
---------	------	---------

Sec 3.8.2.4

- Analysis time windows for EORTC PRO questionnaire added to take into account the unscheduled assessments.
- Repeated measures mixed effect model removed only Time to deterioration analyses retained
- Censoring definitions of time to deteriorations aligned with program standards.

Sec 3.8.3

- Gap Analysis - Gap time only calculated from time to censoring to analysis cutoff date

Sec 3.9.1.4

- Reporting of AEs updated to align with current standards.

Sec 3.9.4

- Reporting of deaths updated to reflect the new death reporting standards,
- Added - most frequent SAE, suspected of study treatment relationship by primary SOC and PT.

Sec 3.9.2

- Updated the definition of baseline for labs for duplicate assessments for CTC gradable and non-CTC gradable labs
- Liver enzyme elevation summary updated to align with the current program standards.

Sec 3.9.4

- Updated clinical notable criteria in vital signs for weight

Sec 3.9.5

- ECOG - analysis window added to take into account the unscheduled assessments
- Analysis for Time to definitive deterioration of ECOG PS added.

Sec 3.10 Subgroup Analyses

- Only median PFS with Kaplan Meier estimates will be presented. No forest-plots will be provided.

Sec 3.11

- Non Compartmental analysis for for PK parameter based on full PK sampling added.
 - Population PK analysis based on Sparse PK samples added
 - Exposure - response analysis for PFS based on Trough PK sampling added.
-

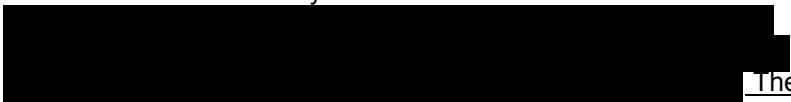
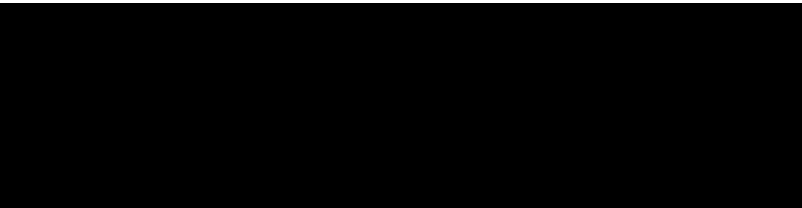
Version	Date	Changes
		<p><i>Sec 3.12</i></p> <ul style="list-style-type: none"> - Archival tumor based Pi3K activation definition revised and now based only on PIK3CA mutation and PTEN loss of expression
V 3.0		<p><u>This Amendment is necessitated to detail the planned analyses for final OS analyses. There is no change proposed to the analyses plan and outputs from the RAP version dated: 21Oct2015 detailed for the Primary CSR. In addition, sensitivity analyses of PFS/ORR/DoR based on Central radiology assessments will be summarized based on data cut-off date:31Aug2015 used for final PFS analyses reporting effort. DoR will also be summarized based on data cutoff for OS analysis to include additional information.</u></p> <p> The following additional analyses have been added for Efficacy:</p> <ul style="list-style-type: none"> - PFS summary and HR as per central radiology assessment to be included. - Comparison of PFS events type and timing of PFS events between local investigator assessment and central radiology assessment - Comparison of Best Overall response between local investigator's assessment and central radiology assessment - Additional analyses for Duration of Overall response and Time to Overall response based on Kaplan Meier method is included. <p></p> <p>PK Section 3.11.2 (population PK Analysis) and 3.11.3 (Exposure response Analyses) are deleted because of limited data.</p>

Table of contents

Table of contents	4
1 Introduction	9
2 Definitions and general methodology	9
2.1 Definitions	9
2.1.1 Study drug and study treatment.....	9
2.1.2 Date of first administration of study drug	9
2.1.3 Date of last administration of study drug	9
2.1.4 Date of first administration of study treatment	10
2.1.5 Date of last administration of study treatment	10

2.1.6	Last date of exposure to study drug/treatment	10
2.1.7	Study day	10
2.1.8	Baseline	11
2.1.9	On-treatment assessment/event	11
2.1.10	Last contact date	12
2.1.11	Screening failure	12
2.1.12	Time Units	12
2.1.13	Time/Assessment windows	13
2.1.14	Data included in the analyses	13
2.1	Definitions of analysis sets	13
2.1.1	Screened patients	13
2.1.2	Full analysis set	13
2.1.3	Safety set	13
2.1.4	Per-protocol set	14
2.1.5	Pharmacokinetic analysis set (PAS)	14
2.1.6	Full Sampling Pharmacokinetic analysis set (FPAS)	14
2.2	Protocol Deviations	15
2.3	Concomitant medications with specific impact on the analysis	15
2.4	Implementation of RECIST	15
2.4.1	Disease Progression	15
2.4.2	Best Overall Response (BOR)	15
2.4.3	Change in imaging modality	16
2.4.4	Determination of missing adequate tumor assessments	16
2.4.5	No Baseline tumor assessments	17
3	Subject disposition, background and demographic characteristics	18
3.1	Enrollment status	18
3.2	Subject disposition	18
3.3	Protocol deviations	19
3.4	Analysis sets	19
3.5	Background and demographic characteristics	19
3.5.1	Basic demographic and background data	19
3.5.2	HPV Status	19
3.5.3	PI3K pathway activation	19
3.5.4	Randomization stratification	20
3.5.5	Diagnosis and extent of cancer	20
3.5.6	Medical History	20

3.5.7	Anti-neoplastic therapy	20
3.6	Study treatment	21
3.6.1	Duration of exposure to study drug/treatment	21
3.6.2	Cumulative dose	22
3.6.3	Dose intensity and relative dose intensity	23
3.6.4	Dose reduction, interruption and permanent discontinuation	24
3.7	Prior and Concomitant therapy	26
3.8	Efficacy evaluation	27
3.8.1	Primary efficacy endpoints	27
3.8.2	Secondary efficacy endpoints	31
3.8.3	Follow-up of the study	37
3.9	Safety evaluation	37
3.9.1	Adverse events data	38
3.9.2	Laboratory data	41
3.9.3	Patient self-reported Mood assessments	43
3.9.4	Vital signs	44
3.9.5	Performance Status	45
3.9.6	Cardiac assessments	47
3.10	Subgroup analyses	48
3.10.1	Efficacy	48
3.11	Pharmacokinetic Analysis	49
3.11.1	Non compartmental analysis	49
3.12	Biomarkers	50
3.13	Interim safety review analyses	52
4	General statistical methodology	52
4.1	Baseline comparability	52
4.2	Center pooling	52
4.3	One-sided vs. two-sided test	53
4.4	Time-to-event analyses	53
4.4.1	Analysis of time-to-event data with ties	53
4.4.2	Kaplan-Meier estimates	53
4.4.3	Hazard ratio	53
4.5	Bayesian methodology for the PFS and OS analyses	54
4.6	Confidence intervals for response rate and disease control rate	54
5	References	54

List of abbreviations

AE	adverse event
ATC	Anatomical Therapeutic Chemical
BOR	best overall response
BSA	body surface area
CBR	clinical benefit rate
CI	Confidence Interval
CR	complete response
CRS	case retrieval strategy
CSP	clinical study protocol
CSR	clinical study report
CT	computed tomography
CTC	Common Terminology Criteria
CTCAE	Common Terminology Criteria for Adverse Events
CV	coefficient of variation
DAR	dosage administration record
DI	dose intensity
DMC	Data Monitoring Committee
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report/record form
EOT	end of treatment
FAS	full analysis set
FPAS	full sampling pharmacokinetic analysis set
GAD	generalized anxiety disorder assessment
GPS	global programming & statistical environment
HLGT	High level group term
HR	Hazard ratio
IRT	interactive response technology
MedDRA	Medical Dictionary for Regulatory Activities
MRI	magnetic resonance imaging
NCRNPD	Non-CR/Non-PD
NMQ	Novartis MedDRA queries
ORR	overall response rate
OS	overall survival
PAS	pharmacokinetic analysis set
PD	progressive disease
PD	protocol deviation
PD	pharmacodynamic
PDI	planned dose intensity
PFS	progression-free survival
PHQ	patient health questionnaire
PK	pharmacokinetic
PoC	proof of concept
PR	partial response

PT	preferred term
RAP	report and analysis plan
RDI	relative dose intensity
RECIST	Response Evaluation Criteria In Solid Tumors
SAE	serious adverse event
SD	stable disease
sd	standard deviation
SEC	safety event categories
SMQ	standardized MedDRA queries
SOC	system organ class
SSD	Study specifications document
SPP	Safety Profiling Plan
TA	tumor assessment
TEAE	treatment-emergent adverse event
ULN	upper limit of normal
UNK	unknown
WHO	World Health Organization

1 Introduction

This document describes the detailed statistical methodology of the Reporting and Analysis Plan (RAP) of study CBKM120H2201: A Phase II multicenter randomized, double blind, placebo controlled study assessing the efficacy of buparlisib (BKM120) plus paclitaxel vs. placebo plus paclitaxel in patients with platinum pre-treated recurrent or metastatic head and neck squamous cell carcinoma.

This RAP Amendment 2 (v3.0) details planned analyses for final OS analysis reporting effort corresponding to protocol version dated: 2nd July 2015. There is no change proposed to the analysis plan, methodology and outputs detailed in RAP Amendment 1(v 2.0) for primary CSR.

In addition, sensitivity analyses of PFS, ORR, and DoR based on central radiology assessments will be summarized based on the data cut-off date used for the final PFS analysis reporting effort.

The data will be analyzed by Novartis and/or a designated CRO. It is planned that the data from all centers that participate in this study will be used.

2 Definitions and general methodology

2.1 Definitions

2.1.1 Study drug and study treatment

Study drug refers to BKM120/BKM120-matching placebo and paclitaxel with which the patient is treated during the study. Study drugs may also be referred to as components of study treatment.

Study treatment refers to BKM120 in combination with paclitaxel or BKM120-matching placebo in combination with paclitaxel.

BKM120-matching placebo will be referred to as “placebo” in the remainder of this document.

2.1.2 Date of first administration of study drug

The date of first administration of study drug is derived as the first date when a non-zero dose of study drug (BKM120/placebo or paclitaxel) is administered. For the sake of simplicity, the date of first administration of study drug is referred to as **start date of study drug**. Start date of study drug is defined for each drug which is part of study treatment.

The date of first administration for BKM120/placebo or paclitaxel is recorded on the corresponding “dosage administration record” (DAR) eCRF page.

2.1.3 Date of last administration of study drug

The date of last administration of study drug is defined as the last date when a non-zero dose of study drug is administered. This date will also be referred to as **last date of study drug**. Last date of study drug is defined for each drug which is part of study treatment.

The date of last administration for BKM120/placebo or paclitaxel is recorded on the corresponding “DAR” eCRF page.

Note 1: Last date of study drug exposure may not be the same as the last date of study drug (see [Section 2.1.6](#)).

2.1.4 Date of first administration of study treatment

The date of first administration of study treatment is derived as the first date when a non-zero dose of any component of the study treatment (BKM120/placebo or paclitaxel) is administered. For the sake of simplicity, the date of first administration of study treatment will also be referred to as *start date of study treatment*.

For example: if the 1st dose of BKM120/placebo is taken on 05JAN2011, and 1st dose of paclitaxel, is taken on 03JAN2011, then the date of first administration of study treatment is 03JAN2011.

2.1.5 Date of last administration of study treatment

The date of last administration of study treatment is defined as the last date when a non-zero dose of any component of the study treatment (BKM120/placebo or paclitaxel) is administered. For the sake of simplicity, the date of last administration of study treatment will also be referred to as *last date of study treatment*.

For example: if the last dose of BKM120/placebo is taken on 15APR2011, and the last dose of paclitaxel is taken on 17MAY2011, then the date of last administration of study treatment is on 17MAY2011.

2.1.6 Last date of exposure to study drug/treatment

The study schedule is organized in cycles of 28 days.

BKM120/placebo is administered daily on a continuous once daily dosing schedule; paclitaxel is administered weekly on D1, D8, D15 and D22 of each cycle.

The *last date of exposure to BKM120/placebo* is the date of last administration of BKM120/placebo.

The *last date of exposure to paclitaxel* is calculated as follows:

- Paclitaxel is administered over several regular weekly doses (every 7 days). The last date of exposure is calculated according to the planned dose schedule of the cycle.
- The last date of exposure to paclitaxel is calculated as: (last date of administration of paclitaxel) + (length of time interval - 1) i.e. [last date of paclitaxel + (7-1)].
If the patient died or was lost to follow-up within last date of administration of paclitaxel + 6 days, the last date of exposure to paclitaxel is the date of death or the date of last contact, respectively.

'Date of last administration of study drug' and 'Date of last contact' are defined in [Sections 2.1.3 and 10](#).

The *last date of exposure to study treatment* is derived to be the later of the last dates of exposure to BKM120/placebo and paclitaxel.

2.1.7 Study day

The study day describes the day of the event or assessment date, relative to the reference start date (randomization date or start date of study treatment). The study day will be displayed in the data listings. It is not to be used for numerical computations for example calculating exposure.

The reference start date is designated as Study Day 1. Study Day -1 is the day that precedes Day 1. Study Day 0 is not defined.

The study day will be calculated as:

The date of the event (visit date, onset date of an event, assessment date etc.) minus reference start date + 1 if the event is on or after the reference start date.

The date of the event (visit date, onset date of an event, assessment date etc.) minus reference start date if the event precedes the reference start date.

The reference start date

- **for all safety assessments** (e.g. adverse event onset, laboratory abnormality occurrence, mood disorder assessment, vital sign measurement, dose interruption etc.) will be the start date of study treatment,
- **for all efficacy assessments** (e.g. tumor assessment, death) will be the randomization date,
- **for any non-safety screening assessments or events** such as baseline disease characteristics or medical history (e.g., time since diagnosis of disease) that occurred prior to randomization the reference start date will be the randomization date.

The study day will be displayed in the data listings.

2.1.8 Baseline

Baseline is the result of an investigation describing the “true” uninfluenced state of the patient, defined as the period from the date of signing any informed consent document to the start date of study treatment or the date of randomization. Assessments, specified to be collected post-dose on the first date of treatment are not considered as baseline values.

For **efficacy evaluations**: the last non-missing assessment, including unscheduled assessments on or before the date of randomization is taken as “baseline” value or “baseline” assessment. In the context of the definition of baseline, the efficacy evaluations also include PRO, biomarkers and ECOG performance status.

For **safety evaluations** (i.e. laboratory, mood disorder assessment and vital signs): the last non-missing assessment, including unscheduled assessments on or before the date of start of study treatment is taken as “baseline” value or “baseline” assessment.

For duplicate laboratory measurements taken at the last assessment date on or before the start date of study treatment, the value of lower CTC grade will be considered as the baseline value.

For Mood questionnaires (PHQ-9 and GAD-7), if the questionnaire is fully completed at screening and partially completed at Cycle 1 Day 1, the questionnaire partially completed will be considered as the baseline and then the total score will be missing at baseline in the analyses. If the screening visit, Cycle 1 Day 1 and the treatment start day are identical, then the screening assessment will be considered as baseline.

If patients have no value as defined above, the baseline value will be missing.

2.1.9 On-treatment assessment/event

Safety summary tables and selected summaries of deaths will summarize only on-treatment assessments/events.

An on-treatment adverse event is defined as any adverse event reported in the following time interval (including the lower and upper limits):

- date of first administration of study treatment; date of last administration of study treatment + 30 days

An on-treatment assessment is defined as any assessment performed after the date of first administration of study treatment i.e. assessments performed in the following time interval (including the lower and upper limits):

- date of first administration of study treatment + 1; date of last administration of study treatment + 30 days

Furthermore, assessments collected post-dose on the date of first administration of study treatment are on-treatment assessments.

If the last date of study treatment is missing, on-treatment assessments/events include any assessment/event recorded in the database and which occur after the start date of study treatment.

Data listings will include all assessments/events, flagging those which are not on-treatment assessments/events.

Note: The date of first administration of study treatment and the date of last administration of study treatment are defined in [Sections 2.1.4](#) and [2.1.5](#), respectively.

2.1.10 Last contact date

The last contact date is derived for patients not known to have died at the analysis cut-off date, and is based on the latest complete date among the following:

- - Actual assessment dates (Labs, vital signs, performance status, tumor imaging, EOT completion, PK sample collection dates).
 - Antineoplastic therapies administered after study drug discontinuation.
 - Adverse events dates
 - “Last known date patient alive” collected on the “Survival information” eCRF page.
 - Study treatment start/end date

The last contact date is defined as the latest complete date from the above list or the cut-off date, whichever comes first.

The last contact date is used for censoring of patients in the analysis of overall survival.

2.1.11 Screening failure

Screening failures are patients who have signed informed consent and failed screening criteria in the study. These patients are not randomized.

Patients who are randomized, but never received study treatment are not screening failures.

2.1.12 Time Units

A month length is 30.4375 days (365.25 / 12). If duration is to be reported in months, duration in days is divided by 30.4375. If duration is to be reported in years, duration in days will be divided by 365.25.

2.1.13 Time/Assessment windows

In order to summarize over time, parameters recorded at each visit (i.e. PRO measures, performance status, PK, vital signs), assessments (including unscheduled ones) will be time-slotted.

The way to deal with multiple assessments within a time window is specified for each assessment type in the relevant sections.

Note: Do not have gap between time window. For assessments done at EoT those should be associated to the corresponding visit window.

2.1.14 Data included in the analyses

The final analyses and the periodic safety reviews by DMC will be performed using all data collected in the database up to the data cut-off date. A cut-off date will be defined for each of these analyses and will be specified in the outputs.

Any data collected beyond the cut-off date for an analysis will not be included in the analysis. Only data with an assessment date or event start date (e.g. vital sign assessment date or start date of an adverse event) prior to or on the cut-off date will be included in the analysis. For example, if the cut-off date is 15 June 2010 then an AE starting on 13 June 2010 will be reported, whereas an AE with start date on 17 June 2010 will not be reported.

All events with an event start date either before or on the cut-off date and an event end date after the cut-off date will be reported as “continuing at the cut-off date”. The same rule will be applied to events starting either before or on the cut-off date and not having a documented end date. This approach applies, in particular, to adverse event and concomitant medication reports. For these events, the end date will not be imputed and therefore will appear as missing in listings.

If it is required to impute an end date to be able to perform a specific analysis (e.g. end date after the cut-off date) the cut-off date needs to be imputed as an end date (to allow for calculation of treatment exposure duration and dose intensity for instance). The imputed date will be displayed and flagged in the listings. A detailed description for handling of missing data including imputation rules will be provided in the RAP Module 8 .

2.1 Definitions of analysis sets

2.1.1 Screened patients

Screened patients consists of all patients who signed informed consent and completed the Screening disposition eCRF page.

2.1.2 Full analysis set

The Full analysis set (FAS) includes all patients who were randomized to study treatment. According to the intent-to-treat principle, patients will be analyzed according to the treatment and strata they have been assigned during the randomization procedure. The FAS will be the main population for analyses of patient disposition, demographics and other baseline characteristics. The FAS will be the primary population for the efficacy analyses.

2.1.3 Safety set

The Safety set includes all patients who received at least one dose of study treatment (either paclitaxel or BKM120/placebo) and had at least one post-baseline safety assessment. Patients will be analyzed

according to the study treatment they actually received, which is defined as the treatment the patient received on the first day of study treatment.

A record of death or “no adverse event” constitutes a valid safety assessment.

Data for patients with no post-baseline safety assessment will only be listed in the FAS.

2.1.4 Per-protocol set

The **Per-protocol set** (PPS) consists of all patients in the FAS who do not have any protocol deviations (see [Section 2.2](#)) that could confound the interpretation of the primary analyses conducted on the FAS. Details regarding protocol deviation leading to exclusion from the PPS will be specified in the study specifications document (SSD). In addition to such protocol deviations additional patient classifications may be used exclude patients from PPS. Patients who are randomized but not treated will be excluded from the PPS.

The PPS may be used to perform sensitivity analysis for the primary efficacy endpoint (i.e. PFS) if Proof of concept (PoC) for the primary endpoint of PFS is established (See [Section 3.8.1.3.2](#)).

2.1.5 Pharmacokinetic analysis set (PAS)

The Pharmacokinetic Analysis Set (PAS) will include all patients who received at least one dose of study medication BKM120 and had at least one evaluable post-treatment concentration measurement.

2.1.6 Full Sampling Pharmacokinetic analysis set (FPAS)

The Full Sampling Pharmacokinetic Analysis Set (FPAS) will include all patients in the PAS who :

- received the planned dose of BKM120 i.e. 100 mg daily every day for the last consecutive 7 days (including Cycle 1 Day 15) preceding full PK profile assessment on Cycle 1 Day 15
- did not vomit within 4 hours of BKM120 dosing after Cycle 1 Day 15 administration
- had at least one dose of paclitaxel prior to the collection of the PK sample for the full PK profile
- had an evaluable full PK profile i.e. at least 7 PK samples (including pre-dose, 0.5h, 1h, 1.5h, and 24h post-dose) available on the day of full PK assessment Cycle 1 Day 15

Note : Patients in the FPAS will also be included in the PAS.

Withdrawal of Informed Consent

Any data collected in the clinical database after a subject withdraws informed consent from all further participation in the trial, will not be included in the analysis data sets. The date on which a patient withdraws full consent is recorded in the eCRF.

Note: Patients who withdraw consent from treatment but consent to post-treatment follow up and/or survival follow up are not considered as having withdrawn full consent and no such withdrawal of consent date is recorded in the eCRF. Such post-treatment assessments will be included in the analysis data sets.

Any data that is entered in the clinical database after the date of full consent withdrawal will be excluded from the analysis sets. Only exceptions are the information entered to close a page to reflect this event and happens after the date of withdrawal of consent e.g End of Treatment, End of Post treatment follow-up, DAR. Deaths happening after withdrawal of consent not confirmed by public registry will be excluded from PFS (if it is an event), OS and deaths in safety.

Third party data e.g PK, biomarker etc, collected in the clinical database without having obtained consent for collection will not be included in the analysis data sets. Data from such samples if collected prior to the date of consent withdrawal will be included in the analysis unless specific request is made by the patient not to analyze the sample.

2.2 Protocol Deviations

The protocol deviations leading to exclusion of a patient from the Per-protocol set (major protocol deviations) are defined in the SSD.CSR Reportable protocol deviations will be identified, summarized and listed.

2.3 Concomitant medications with specific impact on the analysis

Concomitant treatment of BKM120/placebo and weak inducers of CYP3A4 is permitted, during a period of time that should be as short as possible (e.g., less than 1 week). Co-administration of BKM120/placebo with strong CYP3A inducers or inhibitors is prohibited.

Concomitant medications that are inducers or inhibitors of CYP3A4 or substrates of CYP450, non enzyme inducing ant-epileptic drugs, bisphosphonates and permitted QT prolonging drugs as defined in the Appendix 2 of the CSP will be identified and classified (review to be performed by a Clinical Pharmacologist) and then tabulated and/or listed in the Clinical Study Report as appropriate.

2.4 Implementation of RECIST

Response and progression evaluation will be performed according to the RECIST 1.1 (as described in detail in Appendix 6 of the CSP). The text below gives more detailed instructions and rules to provide further details needed for programming.

2.4.1 Disease Progression

Only objective progressive disease (PD) per RECIST 1.1 i.e. PD identified using objective tumor assessment method (e.g. CT scan/MRI, photos for skin lesions, etc.) is considered.

In particular, discontinuation due to PD (from the “End of Treatment Phase Completion” eCRF page), without supporting objective evidence (as defined above), will not be considered as PD in the determination of best overall response(BOR) and in the analysis of progression-free survival(PFS).

2.4.2 Best Overall Response (BOR)

The evaluation of BOR will be assessed by RECIST1.1 criteria. The derivation of Best Overall Response in this study will be based on confirmed CR, PR or SD. The definitions and the details on the derivation are given in Appendix 6 of the CSP.

Only tumor assessments performed before the start of any further anti-neoplastic therapies (i.e. any additional secondary anti-neoplastic therapy with the exception of control drug monotherapy) will be considered in the assessment of BOR.

Further, anti-neoplastic therapies will be identified from the data collected on ‘antineoplastic therapy since discontinuation of study treatment – medication’ , ‘antineoplastic therapy – radiotherapy’ and ‘antineoplastic therapy - surgery’.

Continuation of control drug monotherapy as 1st new anti-neoplastic therapy after end of treatment without prior PD and collected in the ‘antineoplastic therapy since discontinuation of study treatment-Medication’ eCRF page, will not be considered as an anti-neoplastic therapy for the assessment of BOR and for PFS analyses.

Since the tumor assessments are performed 4 weeks after randomization and thereafter every 6 weeks, the standard definition of a best overall response evaluation of “stable disease”, “progressive disease” or “unknown” given in the Appendix 6 of the CSP requires an adjustment.

Best overall response for each patient is determined from the sequence of overall (lesion) responses according to the following rules:

- CR = at least two determinations of CR at least 4 weeks apart before progression
- PR = at least two determinations of PR or better at least 4 weeks apart before progression (and not qualifying for a CR)
- SD = at least one SD assessment (or better) \geq 25 days after randomization date (and not qualifying for CR or PR).
- PD = progression \leq 77 days after randomization date (and not qualifying for CR, PR or SD).
- UNK = all other cases (i.e. not qualifying for confirmed CR or PR and without SD after more than 25 days or early progression within the first 77 days)

Patients with BOR ‘unknown’ will be summarized by reason for having unknown status. The following reasons will be used:

- No valid post-baseline assessment
- All post-baseline assessments have overall response UNK
- New anti-neoplastic therapy started before first post-baseline assessment
- SD too early (< 25 days after randomization and not qualifying SD, CR or PR in any following assessment)
- PD too late (for example, > 77 days after randomization and not qualifying for CR, PR and SD)

2.4.3 Change in imaging modality

Per RECIST 1.1, a change in methodology can be defined as either a change in contrast use (e.g. keeping the same technique, like CT, but switching from with to without contrast use or vice-versa, regardless of the justification for the change) or a change in technique (e.g. from CT to MRI, or vice-versa), or a change in any other imaging modality. A change in methodology will result by default in an UNK (unknown) overall lesion response assessment as per Novartis calculated response. However, another response assessment other than the Novartis calculated UNK response may be accepted from the investigator or the central reviewer if a definitive response assessment can be justified based on the available information.

Potential discrepancies between the modality used and overall lesion response (e.g. change in modality but response is different from ‘Unknown’) will be queried during the data validation process.

2.4.4 Determination of missing adequate tumor assessments

The term ‘missing adequate tumor assessment’ is defined as a tumor assessment (TA) not done or tumor assessment with overall lesion response ‘Unknown’. For the sake of simplicity, a ‘missing adequate tumor assessment’ will also be referred to as a ‘missing assessment’.

As described in Table 3-2 in Section 3.2.9 of RECIST1.1 Novartis guidelines, the PFS censoring and event date options depend on the presence and the number of missing tumor assessments. For example, in the primary analysis of PFS, an event occurring after two or more missing assessments is censored at the last adequate tumor assessment.

An exact rule based on two thresholds D1 and D2 is used to determine whether there is no, one or two missing TAs between the last adequate tumor assessment and a PFS event. This rule is based on the interval between the last adequate tumor assessment date and the event date.

If the interval is greater than threshold D2, then the number of missing assessments is 2 or more. If the interval is greater than D1 and less than or equal to D2 then the number of missing assessments is 1. If the interval is less than D1 then there is no missing assessment

In this study tumor assessments are not being done at equal time intervals or with equal time windows. The first scheduled tumor assessment is 4 weeks after randomization with an allowed time window of ± 3 days. Subsequent tumor assessments are scheduled after every 6 weeks with an allowed time window of ± 4 days. Therefore the thresholds D1 and D2 are defined for two scenarios of last adequate tumor assessment date.

Scenario 1 – Baseline tumor assessment is the last adequate tumor assessment

- The threshold D1 is defined as the protocol-specified interval between the first TA and date of randomization plus the protocol-allowed time window around the first TA. Therefore $D1 = 31$ days (4 weeks*7 + 3 days)
- The threshold D2 is defined as the protocol-specified interval between the first TA and the date of randomization and the interval between the first TA and the second TA plus the sum of protocol-allowed time windows around the first and second TAs. Therefore $D2 = 77$ days (10weeks*7 + 3 days + 4 days)

Scenario 2 – There is at least one post baseline adequate tumor assessment

- The threshold D1 is defined as the protocol-specified interval between the TAs subsequent to the first 1st TA plus twice the protocol-allowed time window around the assessments. Therefore $D1 = 46$ days (6 weeks*7 + 4 days)
- The threshold D2 is defined as twice the protocol-specified interval between the TAs subsequent to the first 1st TA plus twice the protocol-allowed time window around the assessments. Therefore $D2 = 92$ days (12 weeks*7 + 4 days*2)

Therefore, using the D2 definition above, the censoring of an event occurring after ≥ 2 missing TAs (in primary PFS analysis in [Section 3.8.1.2](#)) can be refined as follows: if the distance between the last adequate TA date and the PFS event date is larger than D2 then the patient will be censored and the censoring reason will be 'Event documented after two or more missing tumor assessments'.

Number of patients with one or more than one missing/unknown tumour assessments will be summarized by treatment groups. Also number of patients with one or more than one missing/unknown assessments between last adequate tumour assessment(as per RECIST 1.1) and PFS event/cut-off will also be summarized by treatment group.

2.4.5 No Baseline tumor assessments

As specified in [Table 3-2 in Section 3.2.9 of the Appendix 6 of the CSP], since the timing of disease progression cannot be determined for patients with missing Baseline tumor assessment, these patients

are censored in the PFS analysis at the date of randomization. This rule, only applies to the 'progressive disease' component of the PFS assessment.

Patients without any Baseline tumor assessment who die within D2 (77 days) (Section 2.4.4) from date of randomization will be counted as having an event in the primary analysis of PFS. All deaths will be counted in the overall survival analysis regardless of presence or absence of the baseline tumor assessment.

3 Subject disposition, background and demographic characteristics

3.1 Enrollment status

The number (%) of patients screened and randomized will be summarized by country and center, overall and by treatment.

3.2 Subject disposition

The number (%) of **randomized** patients included in the FAS will be presented overall and by treatment group. The number (%) of screened (**not-randomized**) patients and the reasons for not completing screening phase will also be displayed in a table.

The number (%) of patients in the FAS who are still on treatment, who discontinued the study phases (treatment and post-treatment follow-up) and the reason for discontinuation will be presented overall and by treatment group.

The following summaries will be provided (% based on the total number of FAS patients):

- Number (%) of patients who were randomized (based on IRT)
- Number (%) of patients who were randomized but not treated (based on 'DAR' eCRF page not completed)
- Primary reason for not being treated (based on "End of Treatment Phase Completion" eCRF page)
- Number (%) of patients who were treated (based on 'DAR BKM120/placebo' and 'DAR paclitaxel' eCRF pages completed with non-zero dose administered)
- Number (%) of patients who are still on-treatment (based on the 'End of Treatment Phase Completion' eCRF page not completed)
- Number (%) of patients who discontinued the study treatment phase (based on the 'End of Treatment Phase Completion' eCRF page completed)
- Primary reason for study treatment phase discontinuation (based on the 'End of Treatment Phase Completion' eCRF page completed)
- Number (%) of patients who have entered the post-treatment follow-up (based on the 'End of Treatment Phase Completion' eCRF page completed)
- Number (%) of patients who have discontinued from the post-treatment follow-up (based on the 'End of post-treatment follow-up Phase' eCRF page completed)
- Primary reason for discontinuation from the post-treatment follow-up (based on the 'End of post-treatment follow-up Phase' eCRF page completed).

In a separate summary the reasons for patients not completing the screening phase will be presented overall based on “Screening Phase Disposition” eCRF page.

3.3 Protocol deviations

The number (%) of patients in the FAS with any CSR reportable protocol deviation [REDACTED] will be tabulated by deviation category (as specified in the Study Specifications Document) overall and by treatment group.

The protocol deviations leading to exclusion from analysis sets will be tabulated separately by treatment group. All protocol deviations will be listed.

3.4 Analysis sets

The number (%) of patients in each analysis set (defined in [Section 2.1](#)) will be summarized overall and by treatment group and stratum.

3.5 Background and demographic characteristics

The FAS will be used for all Baseline and demographic summaries and listings.

3.5.1 Basic demographic and background data

Descriptive statistics (mean, median, standard deviation, minimum and maximum) will be presented for continuous variables. The number and percentage of patients in each category will be presented for categorical variables. All summaries will be presented overall and by treatment group in the FAS. For categorical variables, the number and percentage of patients with missing data will be provided.

This analysis will include the following: age, gender, race, ethnicity, height, weight, body surface area (BSA), Alcohol History, Smoking History and ECOG performance status at Baseline.

BMI (kg/m^2) will be calculated as $\text{weight}[\text{kg}] / (\text{height}[\text{m}]^2)$ using weight at Baseline. BSA at Baseline will be calculated using Mosteller formula, i.e., $\text{BSA} (\text{m}^2) = \sqrt{(\text{weight}(\text{kg}) \times \text{height}(\text{cm}) / 3600)}$ using weight at Baseline.

3.5.2 HPV Status

The Historical HPV status, method used to assess HPV status and time since date of assessment of HPV status will be listed overall and by treatment group in the FAS.

The HPV status derived from the central lab will be summarized by treatment group and overall.

3.5.3 PI3K pathway activation

The type of PI3K pathway activation based on archival tumor status will be described (PIK3CA mutation, PTEN loss of expression) overall and by treatment group in the FAS. Characteristics of biomarker assessment samples (archival tumor or fresh tumor) will be described.

3.5.4 Randomization stratification

The number (%) of patients in each stratum (number of prior lines of treatment and region of investigator site) based on data obtained from the IRT system will be summarized overall and by treatment group for the FAS. Discrepancies between the stratum at randomization (IRT) and the actual stratum recorded in the clinical database will be summarized.

3.5.5 Diagnosis and extent of cancer

Summary statistics will be tabulated overall and by treatment group for diagnosis and extent of cancer in the FAS. According to the data collected on the eCRF, this analysis will include the following: primary site of cancer, histologic grade, time since initial diagnosis of primary site, disease stage (at initial diagnosis/at study entry), HPV status from Central Lab assessment, time since initial diagnosis to first recurrence/progression, time since most recent relapse/progression, presence/absence of target and non-target lesions at Baseline, disease metastasis summarized by location and number of metastatic sites .

These event times (in months) are defined as (for example):

$[(\text{reference start date}) - (\text{date of initial diagnosis of primary site})] / 30.4375$

3.5.6 Medical History

Medical history will be summarized overall and by treatment group.

Medical history and current medical conditions, including cancer-related conditions and symptoms entered on the *Relevant medical history / current medical conditions* CRF will be summarized overall and by treatment group and listed in the FAS. Separate summaries will be presented for current and historical medical conditions. The summaries will be presented by primary system organ class (SOC) and preferred term (PT).

Medical history/current medical conditions will be coded using Medical Dictionary for Regulatory Activities (MedDRA). The MedDRA dictionary version used for reporting the study will be specified in the CSR. The latest version available at the time of the analyses will be used.

3.5.7 Anti-neoplastic therapy

Anti-neoplastic therapy will be summarized overall and by treatment group.

Prior anti-neoplastic therapy will be listed in three separate listings: medications, radiotherapy and surgery.

The number (%) of patients receiving prior anti-neoplastic therapy, prior anti-neoplastic medication, prior anti-neoplastic surgery and prior anti-neoplastic radiotherapy respectively will be summarized overall and by treatment group in the FAS.

Prior anti-neoplastic medications will be summarized by :

number of therapies,

therapy type (chemotherapy, targeted therapy, and other),

setting (metastatic, therapeutic, neo-adjuvant, adjuvant),

last therapy before study entry and its outcome (type of therapy, setting, best response),

time since last treatment and its outcome (setting, best response and time since last treatment to progression) will be summarized. Last therapy is defined as the last therapy of the last regimen.. Refer to M8 for definitions of last treatment.

A summary table of regimens using ATC class and preferred term will be provided. Prior radiotherapy and prior surgery will be listed.

The medication therapy type of any combination therapy will be classified based on the following order: chemotherapy, biologic therapy, targeted therapy, hormonal therapy. For example, a combination therapy of chemotherapy and hormonal therapy will be classified as chemotherapy.

Number (%) of patients and best response with prior pretreatment using taxanes and/or cetuximab prior to study entry will be summarized by treatment group.

An additional summary of number (%) of patients by prior Lines of anti-neoplastic therapy received in recurrent/metastatic setting and the medications in each line (e.g Chemotherapy , EGFR inhibitor , Chemotherapy + EGFR inhibitor) will be provided. Any drug (monoclonal antibody or a kinase inhibitor) for which EGFR is known to be as a primary target of action will be considered as an EGFR inhibitor(details provided in RAP Module 8).

Anti-neoplastic medications administered since discontinuation of study treatment will be listed and summarized by ATC class, preferred term, overall and by treatment group.

Anti-neoplastic medications will be coded using the WHO Drug Reference List; Anti-neoplastic surgery will be coded using MedDRA dictionary. The MedDRA dictionary and WHO Drug Reference List version used for reporting the study will be specified in the CSR. The latest version available at the time of the analyses will be used.

3.6 Study treatment

The number (%) of patients who received BKM120/placebo starting dose (100 mg daily) and reduced dose of BKM120/placebo due to dose modification (80 mg daily, 100 mg 5 days out of 7 and 80 mg 5 days out of 7) will be summarized by treatment group based on the Safety set.

Similarly, the number (%) of patients who were administered paclitaxel starting dose (80 mg/m²) and reduced dose of paclitaxel due to dose modification (65 mg/m²) will also be summarized in the safety set by treatment group. The “dose received” (in mg/m²) for paclitaxel is considered:

- ‘80 mg/m²’ for a dose received is between 72 mg/m² and 92 mg/ m² (inclusive).
- ‘65 mg/m²’ for a dose received is between 55 mg/m² (inclusive) and less than 72 mg/ m²
- ‘Any other dose’ for a dose received is below 55 mg/m² or above than 92 mg/ m².

Paclitaxel total dose received in mg/m² at a dosing visit is defined as [Total dose administered(mg) at the visit / BSA(m²) measured at the visit]

In addition, the following parameters will be listed, and summarized by treatment group for the Safety set:

3.6.1 Duration of exposure to study drug/treatment

Duration of exposure to study drug (for BKM120/placebo and paclitaxel) is defined according to dosing regimen for each study drug as outlined in [Section 2.1.6](#).

Duration of exposure (days) = (last date of exposure to study drug) – (date of first administration of study drug) + 1

Duration of exposure to study treatment is considered by taking into account the duration of exposure to each study drug:

Duration of exposure (days) = (last date of exposure to study treatment) – (date of first administration of study treatment) + 1,

The duration includes the periods of temporary interruption. ‘Date of first administration of study drug/treatment’ and ‘last date of exposure to study drug/treatment’ are defined in [Sections 2.1.2/2.1.4](#) and [2.1.6](#).

Duration of exposure to study drug/treatment will be categorized into time intervals. Time interval units will be months. Summaries (i.e. mean, standard deviation etc.) will be displayed in months.

3.6.2 Cumulative dose

Cumulative dose is defined as the total dose given during the study treatment exposure and will be summarized for each study drug separately. For patients who do not receive any study drug, the cumulative dose will be set to zero.

The cumulative dose for BKM120/Placebo in a given dosing period is derived from the DAR defined according to the type dosing schedule of the drug as follows:

- a. Daily Dosing (qd): Sum of doses of the study drug administered to the patient from the start date to the end date of the study drug administration.
- b. Intermittent dosing (5 out of 7 days) : The study administered at X mg /day and dosing schedule of 5 days on 2 days off :

$$\text{Cumulative dose (mg)} = X(\text{mg}) * 5 * W + X(\text{mg}) * \min(D, 5),$$

where

W = integer part of (intermittent dosing period on X(mg)) which represents the number of weeks the patient is on this dosing period prior to the last week of this dosing period, and

D = (duration of the intermittent dosing period on X(mg) – 7*W), which is the number of days the patient is on this X(mg) intermittent dosing in the last week.

Below are some examples of cumulative dose calculations:

Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Planned	Actual		
																Cum Dose	Cum Dose		
Assumed Dose	100	100	80	80	X	X	100	100	100	100	100	X	X	100	100	DoE			
Data from DAR	100 mg qd		80 mg qd		0 mg		100 mg (5/7)								15	1500	1060		
Assumed Dose	100	100	100	100	100	100	100	X	X	100	100	100	100	100	X				
Data from DAR	100 mg qd				0 mg				100 mg (5/7)							15	1500	1200	
Assumed Dose	100	100	X	X	100	100	100	80	80	80									
Data from DAR	100 mg qd		0 mg		100 mg (5/7)			80 mg (5/7)									10	1000	740

Example 1: A patient's duration of exposure is 15 days where a patient has been treated with 100 mg/day on days 1–2, 80 mg/day on days 3–4, zero dose on days 5–6 and 100 mg/day (intermittent dosing) on days 7–15. For the intermittent dosing, $W=1$ and $D=2$. In this case the planned cumulative dose is $100 \times 15 = 1500$ mg and the actual cumulative dose = $2 \times 100 + 2 \times 80 + [(100 \times 5 \times 1) + (100 \times \min(2,5))] = 1060$ mg.

Example 2: A patient's duration of exposure is 15 days where a patient has been treated with 100 mg/day on days 1–7, zero dose on days 8–9 and 100 mg/day (intermittent dosing) on days 10–15. For the intermittent dosing, $W=0$ and $D=6$. In this case the planned cumulative dose is $100 \times 15 = 1500$ mg and the actual cumulative dose = $7 \times 100 + 0 + [(100 \times 5 \times 0) + (100 \times \min(6,5))] = 1200$ mg.

Example 3: A patient's duration of exposure is 10 days where a patient has been treated with 100 mg/day on days 1–2, zero dose on days 3–4, 100 mg/day (intermittent dosing) on days 5–7, 80 mg/day (intermittent dosing) on days 8–10. For the intermittent dosing at each of 80 mg/day and 100mg/day, $W=0$ and $D=3$. In this case the planned cumulative dose is $100 \times 15 = 1500$ mg and the actual cumulative dose = $2 \times 100 + 0 + [(100 \times 5 \times 0) + (100 \times \min(3,5))] + [(80 \times 5 \times 0) + (80 \times \min(3,5))] = 740$ mg.

The cumulative dose for paclitaxel is calculated from the DAR eCRF pages and is expressed in mg/m^2 : It is the sum of ('dose administered' (mg) / BSA (m^2)) during the exposure to paclitaxel.

Notes:

1. $\text{BSA} (\text{m}^2) = \sqrt{(\text{wt}(\text{kg}) \times \text{ht}(\text{cm}) / 3600)}$ (Mosteller formula)
2. Weight in BSA at administration time point is the last non-missing assessment, including unscheduled assessments, on or before the start date of paclitaxel administration (from Vital Signs eCRF page).

For Paclitaxel, which is administered and dosed weekly, cumulative dose will be based on the entire duration of exposure for the cycle.

Cumulative dose should be reported in the same units as the prescribed dose.

3.6.3 Dose intensity and relative dose intensity

Dose intensity (DI) and Relative dose intensity (RDI) will be summarized separately for each study drug, using the individual exposure duration for each study drug.

Dose intensity (DI) for patients who receive study drug (i.e. patients with non-zero duration of exposure) is defined as follows:

$\text{DI} (\text{dosing unit} / \text{unit of time}) = \text{Cumulative dose (dosing unit)} / \text{Duration of exposure (unit of time)}$.

For patients who do not receive any study drug, the DI will be set to zero.

Planned dose intensity (PDI) is the assigned dose by unit of time planned to be given to patients as per protocol in the same dose unit and unit of time as that of the Dose Intensity.

For BKM120/placebo:

- $\text{DI} (\text{mg}/\text{day}) = \text{Cumulative dose (mg)} / \text{duration of exposure (days)}$
- PDI is 100 mg/day
- $\text{RDI} (\%) = \text{DI} (\text{mg}/\text{day}) / \text{PDI} (\text{mg}/\text{day}) \times 100$

For paclitaxel:

- $DI \text{ (mg/m}^2\text{/week)} = \text{Cumulative dose (mg/m}^2\text{)} / (\text{duration of exposure (days)}/7)$.
- PDI is 80 mg/m²/week
- $RDI \text{ (\%)} = DI \text{ (mg/m}^2\text{/week)} / PDI \text{ (mg/m}^2\text{/week)} * 100$

The categorical summaries of RDI and the continuous summaries of RDI (i.e. mean, standard deviation etc.) will be presented.

3.6.4 Dose reduction, interruption and permanent discontinuation

The number (%) of patients with dose reductions or interruptions and permanent discontinuations, and associated reasons, will be summarized separately for each study drug.

‘Dose administered (mg) and dosing frequency from the DAR eCRF pages will be used to determine the dose reductions. Fields with zero dose administered will be used to determine interruptions.

‘Dose permanently discontinued’ ticked box from the DAR eCRF page will be used to determine the permanent discontinuations.

3.6.4.1 Dose interruption

For BKM120/placebo, dose administered at 100 mg daily or 80 mg daily, an interruption is defined as a zero dose on one or more days between two non-zero doses. For BKM120 dose administered at 100 mg 5 days out of 7 and 80 mg 5 days out of 7, an interruption is defined as a zero dose on the scheduled 5 days. As per protocol, a change from continuous schedule to intermittent (5 days out of 7) must be preceded by 2 days without treatment. This protocol required interruption will be counted as an interruption with the reason “As per protocol”.

For Paclitaxel, an interruption is defined as a zero dose administered on the scheduled day of administration.

Note: The last zero dose of BKM120/placebo or paclitaxel (followed by permanent discontinuation) is not considered as a dose interruption. Additionally, two consecutive zero doses of BKM120 (e.g. in the sequence 100 mg daily, 0 mg, 0 mg, 100 mg daily) or paclitaxel will be counted as 1 interruption if the reasons for these two consecutive dose interruption are the same.

The number (%) of dose interruptions along with reasons will be summarized. In addition, reasons for permanent discontinuation from the study drug will be summarized for both BKM120/placebo and paclitaxel.

3.6.4.2 Dose reduction

For BKM120/placebo: A dose reduction is defined as a decrease in dose from the protocol planned starting dose (from 100 mg daily to 80 mg daily) or a change in dosing frequency (i.e. from 80 mg daily to 100 mg 5 days out of 7 or from 100 mg 5 days out of 7 to 80 mg 5 days out of 7), even if the dose decrease or change in dosing frequency has been directly preceded by an interruption. For example, in the sequence 100mg – 0mg – 80mg, the 80mg dose is a dose reduction and 100 mg 5 days out of 7 is the dose reduction in the 80 mg-0mg-100mg 5 days out of 7.

If, due to a dosing error, a patient receives higher than protocol planned starting dose and moves down to the planned starting dose then this is not a dose reduction, however if the change is directly from a higher than planned starting dose down to a lower than protocol planned starting dose, then this is a dose reduction.

If, due to a dosing error, a patient receives lower than previous non-zero dose and resumes later at the protocol specified dose reduction, then lower dose received due to dosing error and protocol specified dose reduction are dose reductions (e.g. in the sequence 100mg od - 70mg od - 80mg od, then 70 mg and 80 mg are dose reductions).

If, due to a dosing error, a patient receives lower than previous non-zero dose and resumes later at lower than previous non-zero dose, then 2 dose reductions will be counted (e.g. in the sequence 100 mg od - 70mg od - 60mg od, 70 mg and 60 mg are dose reductions).

For example,

Sequence	Reduction
<i>With dose change</i>	
100 mg od - 80 mg od-0mg-80 mg od	1 reduction (the 1 st 80mg od)
100 mg od – 100 mg od – 0 mg - 80 mg od	1 reduction (80mg od)
100mg od – 0mg – 80mg od	1 reduction (80mg od)
<i>With dose frequency change</i>	
100mg od - 80mg od - 80mg 5 days out of 7 - 80mg od	2 reductions (80 mg od, 80mg 5 days out of 7)
100 mg od - 100mg 5 days out of 7-0mg-100mg 5 days out of 7	1 reduction (the 1 st 100mg 5 days out of 7)
<i>With dose change & dose frequency change</i>	
100 mg od- 0mg - 80 mg od-0mg-100mg 5 days out of 7	2 reduction (80mg od, 100 mg 5 days out of 7)
100 mg od – 0mg - 80 mg od-0mg-100mg 5 days out of 7-0mg-80mg 5 days out of 7	3 reductions (80 mg od, 100 mg 5 days out of 7, 80 mg 5 days out of 7)
100mg od- 0mg – 80mg od – 0mg– 100 mg 5 days out of 7-0mg- 80mg 5 days out of 7 – 70mg* 5 days out of 7 – 0mg - 80mg 5 days out of 7	4 reductions (80mg od, 100 mg 5 days out of 7, the 1 st 80mg 5 days out of 7, 70 mg 5 days out of 7)
<i>With interruption</i>	
100 mg od-0mg-100 mg od	0 reduction
100mg 5 days out of 7-0mg-100mg 5 days out of 7	0 reduction
<i>With dosing error</i>	
100mg od - 80mg od - 70mg od* - 60mg od	3 reductions (80 mg od, 70mg od and 60mg od)
100mg od - 70mg od* - 100mg od	1 reduction (70mg od)
100mg od - 70mg od* - 80mg od	2 reductions (70mg od, 80mg od)
100mg od - 300mg od* - 250mg od*	0 reduction

Sequence	Reduction
100mg od - 200mg od* - 100mg od*	0 reduction
100mg od - 50mg od* - 100mg od	1 reduction (50 mg od)
<i>With dosing error at the 1st administration</i>	
50mg od* - 100mg od	1 reduction (50 mg od)
50mg od* - 0mg – 50 mg*- 100mg od	1 reduction (50 mg od)
50mg od* - 100mg od – 0mg - 80 mg od	2 reductions (50 mg od and 80mg od)

*dosing error

The number (%) of dose reductions along with reasons will be summarized.

For paclitaxel: A dose reduction is defined as a decrease in dose administered (mg/m²) from the protocol planned dose of 80 mg/m². The count of dose reduction will follow the same logic as BKM120/Placebo. If the patient receives 65 mg/m² dose of paclitaxel, it will be considered as 1 reduction. If the patient further goes on to receive a dose lower than 65 mg/m², it will be considered as 2 reductions. The range of dose for paclitaxel considered for 80 mg/m² and 65 mg/m² is described earlier in Section 3.6.

The number (%) of dose reductions along with reasons will be summarized.

3.6.4.3 Permanent Discontinuation

“Dose permanently discontinued” ticked box from the DAR eCRF page will be used to determine the permanent discontinuations. Reasons for permanent discontinuation from the study drug will be summarized for the study drug.

3.7 Prior and Concomitant therapy

Concomitant therapy is defined as all interventions (therapeutic treatments and procedures) other than the study treatment administered to a patient coinciding with the study treatment period. Concomitant therapy include medications (other than study drugs) starting on or after the start date of study treatment or medications starting prior to the start date of study treatment and continuing after the start date of study treatment.

Concomitant therapy will be summarized for the Safety set.

Concomitant medications will be coded using the WHO Drug Reference List and summarized by ATC class and preferred term. Surgical and medical procedures will be coded using MedDRA dictionary and summarized by primary SOC and PT.

The MedDRA dictionary and WHO Drug Reference List version used for reporting the study will be specified in the CSR. The latest version available at the time of the analyses will be used.

Concomitant medications, procedures and significant non-drug therapies taken concurrently with study treatment will be listed and summarized by ATC class/SOC using frequency counts and percentages. Any prior medications, procedures or significant non-drug therapies starting and ending prior to the start date of study treatment will be listed. Analysis will be based on Safety set.

Analysis will be based on Safety set.

3.8 Efficacy evaluation

Efficacy analyses will include all data observed in patients from FAS population.

The efficacy endpoints based on the tumor assessments will be derived according to the RECIST 1.1 (see Appendix 6 of the CSP for details).

The tumor endpoints derivation is based on the sequence of overall lesion responses at each assessment/time point. However, the overall lesion response at a given assessment/time point will be provided from different sources as illustrated in [Table 3-1](#).

Table 3-1 Sources for overall lesion response

Source 1	Local radiologist/investigator reported overall lesion response
Source 2	Central radiology (from Bioclinica) reported overall lesion response

The primary efficacy analysis will be based on local (investigational) radiologist/Investigator tumor assessment data and will use Source 1 in [Table 3-1](#) as a basis for endpoint derivation. In particular, the final local investigator visit response for each assessment/time point collected on the RECIST 1.1 overall response category collection eCRF page will be used to derive the primary efficacy endpoints. Sensitivity analysis will be based on source 2. The assessment time point dates will be derived by Novartis using the dates of the individual lesion assessments.

3.8.1 Primary efficacy endpoints

The primary efficacy endpoints will be analyzed based on the data observed in the FAS, according to the treatment group patients were randomized and the strata (Number of prior lines of therapy, Region of Investigator Site) they were assigned at randomization.

Progression Free Survival (PFS) is the primary endpoint of this study. The primary analysis will be done after approximately 120 PFS events are documented in the clinical database. The derivation of tumor endpoints is based on the sequence of overall lesion responses at each assessment. The primary and secondary efficacy analyses will be based on the local investigator's tumor assessment data as a basis for endpoint derivation. In particular, the final local investigator visit response for each assessment collected on the RECIST 1.1 overall response category collection eCRF page will be used to derive the primary efficacy endpoints. The assessment time point dates will be derived by Novartis using the dates of the individual lesion assessments

3.8.1.1 Primary variable

PFS derived from investigator assessment will be used as the primary efficacy variable.

The PFS is defined as the time from the date of randomization to the date of the first documented disease progression or death due to any cause. The date of progression is the earliest time when any RECIST progression event is observed with no more than one prior missing assessment. If a patient has not progressed or died at the analysis cut-off date or when the patient receives further anti-

neoplastic therapy, PFS will be censored on the date of the last adequate tumor assessment before the earlier of the cut-off date or start of the further anti-neoplastic therapy date. The PFS censoring reasons are defined in [Section 3.8.1.2](#).

The date of last adequate tumor assessment is the date of the last tumor assessment with overall lesion response of CR, PR, or SD before an event or a censoring reason occurred. In this case the last tumor evaluation date at that assessment is used. If no post-baseline assessments are available (before an event or a censoring reason occurred) the date of randomization is used.

Further anti-neoplastic therapies will be identified from the data collected on all anti-neoplastic therapy (See [Section 2.4.2](#)).

3.8.1.2 Data analysis

The primary analysis will be a comparison of PFS between the two treatment groups. The hazard ratio (HR) for BKM120 + paclitaxel versus placebo + paclitaxel will be estimated by a stratified Cox proportional hazard model, using the randomization strata (prior lines of treatment: 1 vs 2) and region of investigator's site (North America vs Rest of World). The strata information will be based on the data obtained from IRT that was utilized for randomization.

The efficacy criteria are:

- The estimated hazard ratio is equal or less than 0.67 (i.e. 33% reduction in risk of PFS event with BKM120 compared to the placebo arm)

and

- Posterior probability ($HR < 1$) $> 97.5\%$ (PFS treatment benefit of BKM120 compared to the placebo).

The posterior probability in the second criterion will be derived from the Bayesian posterior distribution of the HR. Assuming an uninformative prior distribution, the distribution of the HR will be updated with all available data from the patients included in the FAS.

The distribution of PFS will be estimated using the Kaplan-Meier (K-M) method. The median PFS along with two-sided 95% confidence intervals (CIs) will be presented by treatment group ([Brookmeyer and Crowley 1982](#)).

Additionally, the 25% and 75% percentiles will also be provided. The results will be plotted graphically (K-M plots) by treatment group.

The PFS probabilities at 2, 4, 6 and 8 months, and the associated 95% CIs (Greenwood formula) will be summarized by treatment group.

The median time to censoring will be estimated as a measure of lost to follow up using the KM method having the event and censor indicators reversed. (Schemper, M.; Smith, TL, 1996)

PFS events will be described according to the type of events (death, objective tumor progression) by treatment group and overall.

In addition, a summary of reasons for PFS censoring will be provided by treatment group.

The following categories will be used as appropriate (based on "End of Treatment Phase Completion", "End of post-treatment follow-up Phase", "Withdrawal of informed consent", "Death", "Anti-neoplastic therapy since discontinuation of study treatment-medication" eCRF pages and based on the distance D2 defined in [Section 2.4.4](#)):

- Ongoing without event
- Lost to follow-up
- Withdrew consent
- Adequate assessment no longer available
- New cancer therapy added.

If the distance between the last adequate tumor assessment date and the first of the following dates:

1. Analysis cut-off date
2. Start date of further anti-neoplastic therapy
3. Date of study treatment discontinuation due to consent withdrawal
4. Date of study treatment discontinuation due to loss to follow-up

is smaller or equal to D2 then the censoring reason will be

1. 'Ongoing',
2. 'New cancer therapy added',
3. 'Withdrew consent'
4. 'Lost to follow-up', respectively.

However, if this distance is larger than D2 and there is no PFS event (see [Section 2.4.4](#)), then the censoring reason will always default to 'Adequate assessment no longer available'. If the distance between the last adequate tumor assessment date and the PFS event date is larger than D2 then the patient will be censored and the censoring reason will be 'Event documented after two or more missing tumor assessments'.

3.8.1.3 Sensitivity and other supportive analyses of the primary endpoint PFS

Some sensitivity PFS analyses may be performed for the full population. No other inferential statistics will be provided.

As a supportive evidence, the primary PFS analyses performed based on the data cut-off date: 31 Aug 2015 used for primary PFS analysis reporting effort will be repeated using data from central radiology assessments (Source 2 in [Table 3-1](#)) based on patients in FAS.

For central radiology, data from the two independent central readers will be listed together with the final (i.e. either of the 2 readers data if they agree or the adjudicated data between the two central readers if the 2 readers disagree) data. Differences in overall responses between local radiology (Source 1, [Table 3-1](#)) and central radiology (Source 2, [Table 3-1](#)) will be summarized.

The following sensitivity PFS analyses will be performed for the patients in FAS:

- Comparison of PFS events type and timing of PFS event between local investigator's assessment and central radiology assessment
- Summary of the discordance on PFS events type and timing of PFS between local investigator's assessment and central radiology assessment
- Summary of missing/unknown tumor assessments as per central radiology review (Section 3.8.1.3.3)

3.8.1.3.1 Stratification as per eCRF

The Primary Analyses will be repeated using the stratification variables as per eCRF instead of IRT if the efficacy criteria for primary analyses using the stratification factors based on IRT given in Section 3.8.1.2 are met..

Patients with Prior Lines of therapy at recurrent/metastatic setting =1 and region = North America, will be considered Strata 1. Patients with Prior Lines of therapy at recurrent/metastatic setting =1 and region = Rest of the world, will be considered Strata 3. Patients with Prior Lines of therapy at recurrent/metastatic setting =2 and region = North America, will be considered Strata 2. Patients with Prior Lines of therapy at recurrent/metastatic setting =2 and region = Rest of the world, will be considered Strata 4.

3.8.1.3.2 Per-protocol set analysis

If the PoC is established for the primary endpoint of PFS, the primary analyses will be repeated in the PPS.

3.8.1.3.3 Missing/unknown tumor assessments (TAs)

The purpose of the below analyses is to gain an insight as to whether the TAs have been carried out in accordance with the protocol and to understand if any meaningful discrepancies exist between the pattern of missing/unknown assessments between local and central reviews and between treatment arms with each of the two sources.

The following time windows (in weeks) will be constructed: [0, 7], [7, 13], [13, 19], etc. for each patient, where '0' is the patient's date of randomization. Every time-window (with the exception of the initial, broader one) is centered at the scheduled time of TA, i.e., around week 10, week 16 for second and third window, respectively, etc. A patient is considered 'at risk' of missing their TA for any one of these time-windows if the patient either:

- was 'on study' for at least the first 2 weeks of the time-window (4 weeks for the first time window), i.e., if the patient is ongoing at the time of the scheduled TA,
- or
- discontinued treatment due to documented disease progression within the specific time window.

For example, if a patient discontinued during Week 15, then she would have been 'at risk' of a missing/unknown TA for the [13, 19] week time-window. For the purpose of this analysis, 'unknown' TAs (i.e., evaluations with an overall lesion response of 'unknown') will be considered to be missing. However, a clear distinction between 'truly missing' and 'present but unknown' needs to be done in the derived dataset to allow for both combined, i.e. missing and unknown treated the same, and separate analyses.

TAs performed after a documented disease progression will not be considered. In other words, the final time-window for which a patient would be at risk of a missing/unknown scan would be that during which the documented progression occurred.

For patients without documented progression, all TAs are considered up to the earliest of the following dates: death, the analysis cut-off, start of new antineoplastic therapy, discontinuation for either withdrawal of consent or loss to follow-up.

The results will be presented by treatment arm and separately for investigator/local and central radiology using the same cut-off date. Number of patients with at least one missing/unknown TA will be presented together with the following categories: number of patients with 1, 2, 3, 4, 5, >5 missing/unknown TAs.

Since the two treatment arms are expected to differ in the follow-up “tails” the following summaries will be displayed in addition: number of patients with at least one missing/unknown TA between the last adequate TA and PFS event or cut-off, number of patients with at least one missing TA after the study treatment discontinuation.

3.8.2 Secondary efficacy endpoints

All secondary efficacy analyses will be reported by treatment group based on the local investigator’s assessment of the tumor assessment data. Analyses will be based on the FAS unless otherwise specified.

Sensitivity analyses for tumor related endpoints will also be performed based on central radiology assessments.

3.8.2.1 Key Secondary efficacy endpoints

The key secondary objective is to determine whether the treatment BKM120 plus paclitaxel provides an improvement in overall survival compared to treatment with placebo plus paclitaxel. The key secondary analysis will be done after approximately 112 deaths are documented in the clinical database.

Overall survival (OS) is defined as the time from date of randomization to date of death due to any cause. If a patient is not known to have died by the date of analysis cut-off, OS will be censored at the date of last contact (see [Section 2.1.10](#)).

All deaths occurring before the cut-off date will be counted in the overall survival analysis regardless of presence or absence of the Baseline tumor assessment. Data collected in the ‘Death’ eCRF page will be used.

The analyses for OS will be based on the FAS.

OS will be compared between the two treatment groups. The HR for BKM120 + paclitaxel versus placebo + paclitaxel will be estimated by a stratified Cox proportional hazard model, using the randomization strata (i.e. prior lines of therapy and region (North America vs Rest of the world)). HR will be presented together with the 95% confidence interval.

In order to guide the Phase III design for this indication, Proof of Concept (PoC) for Overall Survival is defined based on the following criteria:

- The estimated hazard ratio is equal or less than 0.77 (i.e. 23% reduction in risk of OS event with BKM120 compared to the comparator arm)

and

- Posterior probability ($HR < 1$) > 90% (OS treatment benefit of BKM120 compared to the control).

The posterior probability of $HR < 1$ will be derived from the Bayesian posterior distribution of the HR for Overall Survival. Assuming an uninformative prior distribution, the distribution of the HR will be updated with all available data from the patients included in the FAS. The above proof of concept analysis for OS will only be done after approximately 112 deaths are documented in the clinical

database. All the other analyses of OS as described below may be done at the time of the Primary analysis of the study. The PoC for OS will be assessed irrespective of PFS PoC outcome.

The distribution of OS will be summarized using the K-M method. The median OS along with two-sided 95% CIs will be presented by treatment group. Additionally, the 25% and 75% percentiles will also be provided. The results will also be plotted graphically (K-M plots) by treatment group.

The proportion of patients alive at 3.5, 7, 9, 11 and 13 months and the associated 95% CIs will be summarized by treatment group.

Number of deaths as well as the primary reason for death (study indication, other) will be displayed by treatment group. A listing of patients displaying the reason of death entered by the investigator will be provided.

In addition, censoring reasons ('Alive' or 'Lost to follow-up') will be described by treatment group.

'Lost-to-follow-up' will be determined using the "End of Treatment Phase Completion", "End of post-treatment follow-up Phase", and "Survival information" eCRF pages. Patients not known to have died will be censored for 'Lost to follow-up' if the time between their last contact date and the analysis cut-off date is longer than 3 months and 2 weeks (104 days).

3.8.2.2 Overall response rate (ORR)

Overall response rate (ORR) is defined as the proportion of patients with confirmed best overall response of complete response (CR) or partial response (PR) based on the local investigator's assessment according to RECIST 1.1 (see Appendix 6 of the CSP). However, patients with non-measurable lesions only at Baseline will be included as 'responders' if they achieve a complete response.

Proportions of patients with ORR will be presented by treatment group along with exact 95% CIs (Clopper & Pearson, 1934).

ORR will also be summarized using the data from central radiology assessments based on the data cut-off date used for primary PFS analysis reporting effort (Source 2 in [Table 3-1](#)).

3.8.2.3 Duration of Overall Response and Time to Overall Response

Duration of Overall response (DoR) applies only to FAS responders, i.e., patients whose best overall response was CR or PR. This analysis will include only responders and therefore it might introduce a bias.

DoR is defined as the elapsed time between the date of first documented response (CR or PR, which has to be confirmed subsequently) based on the local investigator's assessment according to RECIST 1.1 (see Appendix 6 of the CSP) and the following date of event defined as the first radiologically documented disease progression or death due to study indication.

DoR will be summarized for patients with a confirmed CR or PR using descriptive statistics. The aspects of censoring as discussed for PFS will also apply here.

DoR will also be summarized for both local investigator and central radiology assessment based on the data cut-off date to be used for the final OS analysis reporting effort (Source 2 in [Table 3-1](#)).

Discordance on Best Overall Response between local investigator's assessment and central radiology assessment will be summarized by treatment arms based on the data cut-off date used for primary PFS analysis reporting effort.

Time to overall response (TTR) applies only to FAS responders, i.e., patients whose best overall response was CR or PR. This analysis will include only responders and therefore it might introduce a bias.

TTR is defined as the time from the date of randomization to the date of first documented response (CR or PR, which has to be confirmed subsequently) based on the local investigator's assessment according to RECIST 1.1 (see Appendix 6 of the CSP). The date of event is defined as the date the response is first determined and not using the date the response is confirmed. Descriptive summary statistics will be reported on the confirmed responders subset only.

Additionally, Kaplan-Meier plot for Duration of response and TTR based on local investigator assessment and central radiology assessment will be provided by treatment arms.

TTR will also be summarized using the data from central radiology assessments based on the data cut-off date: used for primary PFS analysis reporting effort (Source 2 in [Table 3-1](#)).

Disease control rate (DCR) is defined as the proportion of patients with a best overall response of CR, PR or stable disease (SD), based on the investigator assessment (see RECIST Version 1.1 guidelines, Appendix 6 of CSP).

Proportions of patients with DCR will be presented by treatment group along with exact 95% CIs ([Clopper & Pearson, 1934](#)). DCR will also be summarized using the data from central radiology assessments based on the data cut-off date: used for primary PFS analysis reporting effort (Source 2 in [Table 3-1](#)).

At the time of Primary Analysis, many of the patients who responded were getting censored for duration of response. Thus for OS Analysis, Duration of response summaries based on local investigator assessment as well as central radiology assessment is based on OS Cut-off to present more mature data, with reduced number of censoring.

3.8.2.4 Patient Reported Outcomes – Health Related Quality of Life

The EORTC QLQ-C30 questionnaire along with the head and neck cancer module (QLQ-HN35) will be used to collect data on patient's QoL.

The global health status/QoL scale and pain score of the QLQ-C30 and QLQ-HN35, respectively, are identified as the primary patient-reported outcome (PRO) scales of interest. Physical functioning, emotional functioning and social functioning scale scores of the QLQ-C30 and the head and neck cancer symptom scales for swallowing, sense and speech of the QLQ-HN35 are identified as secondary patient-reported HRQoL variables of interest. The FAS population will be used for all PRO summaries of interest.

Scoring and handling of missing QoL data from the EORTC QLQ-C30 and QLQ-HN35 will be in accordance to the EORTC scoring manual (Fayers 2001). No imputation will be applied if the total or subscale scores are missing at a visit. For each domain and item ([Table 3-2](#) and [Table 3-3](#)), a linear transformation is applied to standardize the score between 0 and 100. Higher score values for a functional scale and global health status/QoL indicate higher functioning and health-related quality of life, respectively. Higher score values for symptoms scales indicate greater symptomatology or problems.

Table 3-2 EORTC QLQ-C30 scales

Domain	(Sub)Scale	Item numbers
Global health status/QOL	Global health status/QOL	29, 30
Functional scales	Physical functioning	1 to 5
	Role functioning	6, 7
	Emotional functioning	21 to 24
	Cognitive functioning	20, 25
	Social functioning	26, 27
Symptom scales	Fatigue	10, 12, 18
	Nausea and vomiting	14, 15
	Pain	9, 19
	Dyspnea	8
	Insomnia	11
	Appetite loss	13
	Constipation	16
	Diarrhea	17
	Financial difficulties	28

Table 3-3 EORTC QLQ-HN35 scales

Domain	(Sub)Scale	Item numbers
Symptom scales	Pain	31-34
	Swallowing	35-38
	Senses problems	43,44
	Speech problems	46,53,54
	Trouble with social eating	49-52
	Trouble with social contact	48,55-58
	Less sexuality	59,60
	Teeth	39
	Opening mouth	40
	Dry mouth	41
	Sticky saliva	42
	Coughing	45
	Felt ill	47
	Pain killers	61
	Nutritional supplements	62
	Feeding tube	63
Weight loss	64	
Weight gain	65	

Patient-reported HRQoL questionnaires (EORTC QLQ-C30 and QLQ-HN35) are planned to be administered at screening, every 6 weeks after randomization, and at the End of Treatment assessment.

The following time-based intervals (Table 3-4) will be used to group the HRQOL data over time.

Day x is defined as date of HRQOL administration – randomization date + 1.

Table 3-4 Time based intervals to group HRQOL

Visit	Time Interval
Baseline	Time period on or before the randomization date, i.e., \leq Study Day 1
Week 6	Day 2 to Day 63
Weeks 12, 18 etc. or EOT	+/- 3 weeks centered around the planned assessment i.e. days 64, days 105 for Week 12 (3rd assessment) days 106, days 147 for Week 18 (4th assessment) days etc. (week x -3*7+1 ; week x +3*7)

Time windows are applicable for descriptive summary of PRO data by visit. It is not applicable to time to deterioration analysis since all post-baseline assessments will be considered.

If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. For the baseline time window the target would be study day 1 (i.e., just before the patient starts randomized treatment phase); for the first post-baseline assessment the target would be study day 43 (week 6+1) and thereafter it would be every 6 weeks, corresponding to the planned assessment times of day 84 (week 12), day 126 (week 18), etc. If 2 assessments within a time window are equidistant from the target date then the assessment obtained prior to the target date will be considered. If the closest assessment to the target date has two questionnaires filled out on the same date, then the average of these assessments will be used for each subscale score.

The number of patients completing PRO assessments and the number of patients missing/expected to have completed PRO assessments as reported by the investigational sites will be summarized by each treatment group for each scheduled assessment time point in the FAS.

Other exploratory analysis may be performed to better understand any missing data patterns identified for PRO assessments.

Descriptive statistics (N, mean, median, SD, Q1, Q3, minimum, maximum) will be used to summarize the individual item and scored sub-scale scores of PRO data at each scheduled assessment time point in the FAS for all the sub-scales defined in Table 3-2 and Table 3-3. Additionally, change from baseline in the sub-scale scores at the time of each assessment will be summarized. Patients with an evaluable baseline score and at least one evaluable post baseline score during the treatment period will be included in the change from baseline analyses.

Time to deterioration in the global health status / quality of life scale, physical functioning, emotional functioning, social functioning, and head and neck cancer symptoms scales for pain, swallowing, senses, and speech will be assessed for the two treatment arms. The survival distributions will be presented descriptively using Kaplan-Meier curves for global health status / quality of life scale of EORTC-QLQ-C30 and head and neck cancer symptoms scales for pain, swallowing, senses, and speech. Time to deterioration in the physical functioning, emotional functioning, social functioning subscales of EORTC-QLQ-C30 will only be listed. Summary statistics from the Kaplan-Meier distributions will be determined, including the median time to deterioration and the proportions of patients without deterioration in PRO scales of interest at 3 and 6 months. Both point estimates and 95% confidence intervals will be presented.

Definitive Deterioration in PRO (EORTC-QLQ-C30- global health status / quality of life scale) scales will be defined as a decrease in the subscale score by at least 10% compared to baseline, with no later increase above this threshold observed during the course of the study. A single measure reporting a decrease of at least 10% is considered evaluable only if it is the last observation available for the patient.

Note: Definitive Deterioration in PRO scales (EORTC-QLQ-HN35) will be defined as an increase in the subscale score by at least 10% compared to baseline, with no later decrease above this threshold observed during the course of the study. A single measure reporting an increase of at least 10% is considered evaluable only if it is the last observation available for the patient.

Time to deterioration is the number of days between the date of randomization and the date of the assessment at which definitive deterioration is seen. If a patient has not had an event prior to analysis cut-off or start of another anticancer therapy, time to deterioration will be censored at the date of the last PRO evaluation. Only assessments collected while the patient is on treatment and on or before the end of treatment visit will be included in the PRO time to definitive deterioration. If a definitive deterioration is observed after 2 or more missing assessments, this event will the patient is censored at the date of the last available assessment prior to the deterioration. For example, if a patient has an assessment at week 6, misses the following two assessments on weeks 12 and 18 and a definitive deterioration is observed on week 23, then the patient will be censored at week 6.

Death will be considered as deterioration of symptoms/PRO if it occurs close to the last available assessment, where “close” is defined as twice the planned period between two assessments. This avoids overestimating the time to definitive worsening in patients dying after an irregular assessment scheme. Patients who die after more than twice the planned period between two assessments since the last assessment are censored at the date of their last available questionnaire. For example, if the last assessment is at week 6 and the patient dies on week 16, the deterioration date will be week 16. On the other hand, if the last assessment is at week 6 and the patient dies on week 24, which is after more than twice the planned period between two assessments since the last assessment (week 6), then the patient will be censored at week 6. Patients receiving any further anti-neoplastic therapy (See [Section 2.6.3](#)) before definitive worsening will be censored at the date of their last assessment before starting this therapy. Patients that have not worsened as of the cutoff date for the analysis will be censored at the date of their last assessment before the cutoff.

Patients without baseline PRO or without any post-baseline PRO will be censored at the date of randomization with censoring reason being ‘No baseline score’ or ‘No post-baseline score’, respectively. If the time interval between the last assessment date and the earliest of the following dates is \leq twice planned interval protocol assessments i.e. 98 days ($2*6\text{weeks} + 2*7\text{ days}$)(Table 7-6 of CSP) and there is no definitive deterioration:

1. analysis cut-off date
2. End of Treatment Date
3. start date of further antineoplastic therapy
4. date of consent withdrawal
5. Visit date of study treatment discontinuation due to lost to follow-up

then the censoring reason will be

1. Deterioration-free at the last assessment until cut-off,
2. Deterioration-free at the last assessment until EOT,

3. New anticancer therapy added,
4. Withdrew consent,
5. Lost to follow-up.

If this time interval is larger than 98 days and there is no definitive deterioration event, then the censoring reason will always default to 'Adequate assessment no longer available'. However, if definitive deterioration occurred more than 98 days since the last assessment then the censoring reason will be 'event after ≥ 2 missing assessments'.

Patients receiving any further antineoplastic therapy before definitive worsening will be censored at the date of their last assessment before starting this therapy. Patients that have not worsened as of the cutoff date are censored at the date of their last assessment before the cutoff.

3.8.3 Follow-up of the study

Summary information on the follow-up of patients is displayed in order to describe the maturity of data and quality of follow-up.

Follow-up of the study will be summarized using the following methods to provide a comprehensive assessment of follow-up for patients in each treatment group in the FAS.

Summary of duration between randomization and cut-off date, and follow-up times for PFS/OS, which are defined as follows:

- Randomization (recruitment) period = (Date of last patient randomized - Date of first patient randomized + 1)/30.4375 (months).
- Duration between randomization and data cut-off date = (Cut-off date - Date of randomization + 1)/30.4375 (months).
- Follow-up time = (Date of event or censoring - Date of randomization + 1)/30.4375 (months) regardless of censoring. Date of event or censoring is the same as the one defined for the main analyses.

Gap analysis: Summary of gap time (months) for PFS/OS follow-up as compared to cut-off date. Gap time for PFS/OS is defined as follows:

- For patients who are censored regardless of follow-up status
Gap time = (Analysis cut-off date - censoring date)/30.4375

3.9 Safety evaluation

The assessment of safety will be based mainly on the frequency of on-treatment adverse events and on the number of laboratory values that fall outside of pre-determined ranges. Other safety data (e.g. ECG, mood disorders, and vital signs) will be considered as appropriate.

Risks identified in the Safety Profiling Plan (SPP) will mainly be analysed by frequency of events, further exploratory analyses may be performed as appropriate. The latest version of the SPP available at the time of the analyses will be used.

The safety summary tables include only **on-treatment** assessments (refer to [Section 2.1.9](#) for definition). All safety assessments are listed and those collected outside of the on-treatment window are flagged.

All safety outputs will use the Safety set, unless otherwise specified.

3.9.1 Adverse events data

3.9.1.1 Dictionary Coding of AEs

Adverse events will be coded according to the MedDRA dictionary and will be reported by primary SOC and PT, and by Standardized MedDRA Queries (SMQs) and Novartis MedDRA Queries (NMQs) for safety topics of interest.

The MedDRA dictionary version used for reporting the study will be the version used for coding the study and will be specified in the CSR. The latest version available at the time of the analyses will be used.

3.9.1.2 Grading of AEs

Adverse events will be graded according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.03.

The CTCAE represents a comprehensive grading system for reporting the acute and late effects of cancer treatments. CTCAE grading is by definition a 5-point scale generally corresponding to mild, moderate, severe, life threatening, and death. This grading system inherently places a value on the importance of an event although there is not necessarily proportionality among grades (a grade 2 is not twice as bad as a grade 1).

If CTCAE grading is not available for an adverse event, Grades 1 – 4 corresponding to the severity of mild, moderate, severe, and life-threatening will be used. CTCAE Grade 5 (death) is not be used in this study; rather, death information will be collected in the “death” eCRF page.

3.9.1.3 General rules for AE reporting

Treatment-emergent adverse events (TEAE) are events that occurred on treatment or events present prior to the study treatment start but increased in severity based on preferred term when the patient is on-treatment.

All deaths including those that occurred on-treatment and post treatment, will be listed, post treatment deaths will be flagged. In addition to the on-treatment deaths and additional summary of all deaths (on-treatment + post-treatment) by treatment will be provided.

All AEs summary tables will present only TEAEs. The number (%) of patients reporting any AE will be summarized by primary SOC, PT and treatment group. The most common AEs reported ($\geq 10\%$ in either group for each preferred term) will be presented in descending frequency according to its incidence in BKM120 + paclitaxel treatment group starting from the most common event.

Additional summaries will be provided by maximum toxicity grade and relationship to study treatment.

If a patient reported more than one AE with the same PT, the AE with the maximum toxicity grade will be presented. If a patient reported more than one AE within the same primary SOC, the patient will be counted only once with the maximum toxicity grade at the primary SOC level, where applicable.

Summaries will be provided for study treatment related adverse events, death, serious adverse event, other significant adverse events leading to discontinuation and adverse events leading to dose adjustment. Adverse event tables presented by maximum CTCAE grade will provide the following information: (1) maximum individual grade, (2) maximum G3/G4 and (3) Any grade.

Note: For in-text tables, adverse event tables presented by maximum CTCAE grade will provide the following information: Any grade and maximum G3/G4.

3.9.1.4 AE Summaries

The following AE summaries will be produced by treatment group for all patients in Safety set:

- On-treatment deaths, by PT (*This indicates the number of deaths due to study indication and those due to AE*)
- Deaths by primary SOC, PT. (*All deaths (on-treatment + post treatment) are included*)
- SAEs, regardless of study treatment relationship, by primary SOC, PT and maximum CTCAE grade
- SAEs suspected to be related to the study treatment, by primary SOC, PT and maximum CTCAE grade
- Most frequent SAEs (at least 2%), regardless of study treatment relationship, by primary SOC and PT
- Most frequent SAEs (at least 2%), suspected of study treatment relationship, by primary SOC and PT
- AEs, regardless of study treatment relationship by primary SOC, PT and maximum CTCAE grade
- AEs suspected to be related to the study treatment by primary SOC, PT and maximum CTCAE grade
- Most frequent AEs (at least 10%), regardless of study treatment relationship by PT and maximum CTCAE grade (All grades and Grade 3/4)
- Most frequent AEs (at least 10%), suspected to be related to the study treatment by, PT and maximum CTCAE grade (All grades and Grade 3/4)
- AEs leading to study drug permanent discontinuation, regardless of study treatment relationship, by primary SOC, PT and maximum CTCAE grade
- Most frequent AEs (at least 2%) leading to study drug permanent discontinuation, regardless of study treatment relationship, by primary SOC, PT and maximum CTCAE grade
- AEs leading to study drug dose adjustment, regardless of study treatment relationship, by primary SOC, PT and maximum CTCAE grade
- AEs leading to study drug temporary interruption, regardless of study treatment relationship, by primary SOC, PT and maximum CTCAE grade

- AEs requiring additional therapy, regardless of study treatment relationship, by primary SOC, PT and maximum CTCAE grade
- Non-serious AEs, regardless of study treatment relationship by primary SOC, PT, maximum CTCAE grade *Note: Required for reporting trial results to CT.GOV*

Deaths and other SAEs will also be listed for patients who were screening failures and those who were randomized but not treated.

3.9.1.5 Specific Safety Event Categories

Specific Safety Event Categories (SEC) consist of adverse events for which there is a specific clinical interest in connection with BKM120 treatment (i.e. where BKM120 may influence a common mechanism of action responsible for triggering them) or adverse events which are similar in nature (although not identical).

General considerations for adverse events described in [Section 3.9.1](#) are applicable to SECs. Each of these SECs uses MedDRA terms (SMQ, NMQ, HLTG etc.) to group preferred terms for which there is a specific clinical interest. One SEC can be defined by one or several MedDRA categories. The MedDRA preferred terms used to define the SECs are defined in the program Case Retrieval Strategy (CRS) document and will be summarized. The latest version of the CRS document available at the time of the analyses will be used.

The following SECs will be reported:

- Mood disorders
- Hyperglycemia
- Hypersensitivity, rash (including DRESS, photosensitivity)
- Liver toxicity
- Posterior Reversible Encephalopathy Syndrome(PRES)
- Asthenia, Fatigue
- Nausea, Vomiting, Diarrhea
- Pneumonitis

Additional SECs may be reported if there are any updates to the CRS at the time of the analyses.

Standard data of analysis for all SECs will be conducted as follows:

- The number (%) of patients will be reported by SEC risk, maximum CTCAE grade and by treatment group.
- The number (%) of patients will be reported by SEC risk, by MedDRA category, by preferred term, maximum CTCAE grade and by treatment group.
- All AEs of patients having at least one AE in that SEC will be listed.
- For particular SECs (hyperglycemia, Hypersensitivity/rash, mood disorder for BKM120), depending on the number of events, the time to first occurrence of any grade event (or any grade \geq 2) will be summarized using Kaplan-Meier method. Median time to onset and 95% CI will be summarized. Kaplan-Meier plots will be generated.

The version of MedDRA version and NMQ dictionary date of the Case Retrieval Sheet used for the analyses will be described in a footnote. The Case Retrieval Sheet used for the analyses will be listed.

3.9.1.6 Time to onset of CTCAE Grade 2 or worse SEC

Time to onset of CTCAE Grade 2 or worse SEC is defined as the time from the start of treatment to the start date of the first incidence of CTCAE Grade 2 or worse SEC i.e. time in days is calculated as (start date of first occurrence of Grade 2 or worse SEC) – (date of first dose of study treatment) +1. A patient will be censored for time to onset of Grade 2 or worse occurrence if the patient:

- does not experience any post-baseline Grade 2 or worse SEC
- dies without experiencing Grade 2 or worse SEC
- receives a new anti-neoplastic therapy without experiencing Grade 2 or worse SEC or before the Grade 2 or worse SEC has occurred
- discontinues from the study treatment without experiencing Grade 2 or worse SEC (up to 30 days after last date of study treatment administration)
- is still ongoing at the analysis cut-off date without experiencing Grade 2 or worse SEC,

In the absence of grade 2 or worse SEC, the censoring date will be the earliest date from the following dates: last date of administration of study treatment + 30 days, analysis cut-off date, new anti-neoplastic therapy start date, death date and last contact date.

3.9.2 Laboratory data

3.9.2.1 Grading of laboratory data

Laboratory data grades of severity will be derived according to the Common Terminology Criteria for Adverse Events (CTCAE) v4.03. CTCAE grade is applied to the converted measurement in SI units. A severity grade of 0 will be assigned when the value is within normal limits. In the case when a local laboratory normal range overlaps into the higher (i.e. non-zero) CTCAE grade, the CTCAE grade will be taken.

For laboratory tests for which grades of severity are not defined by CTCAE, results will be graded by the low/normal/high classifications based on laboratory normal ranges.

Laboratory values with missing units or normal range may not be able to be graded or included in laboratory tables.

In cases of multiple laboratory measurements for gradable parameters taken on the same last assessment date:

- The lowest absolute grade should be considered for baseline.
- For duplicate laboratory measurements with same grade but different values taken at the last assessment date on or before the start date of treatment, the baseline will be chosen as the values closest to the ULN if the duplicate values are at a grade > 0 and farthest from ULN if the duplicates are at grade 0. This rule will apply for any lab parameter where increasing value implies increasing grade eg. Total Bilirubin, AST, ALT etc. For labs where decreasing values

imply increasing grade e.g. Platelets, WBC etc, the same rule will be applied with respect to the LLN. Details are provided in RAP M8.

If several records have the same absolute grade, but in different directions, 2 baselines should be created, the record with grade below 0 should be the baseline of the 'Hypo' parameter, and the other record should be the baseline for the 'Hyper' parameters.

For non-gradable labs with duplicate laboratory measurements taken at the last assessment date on or before the start date of study treatment:

- If both within normal range: take the value which is closest to the average of ULN and LLN.
- If one within normal range and the other outside: take the one within normal range.
- If both outside normal range: take the one closest to the normal range.

In some cases these duplicate labs may come from different labs with different LNRs. More details on the derivation of the Baseline in such cases are described in the RAP M8.

Laboratory values with missing units or normal range may not be able to be graded or included in laboratory tables.

A severity grade of 0 will be assigned when the value is within normal limits. In the case when a local/central laboratory normal range overlaps into the higher (i.e. non-zero) CTCAE grade, the CTCAE grade will be taken.

3.9.2.2 Laboratory data summary

The summary of on-treatment laboratory evaluations will be presented for two groups of laboratory tests (hematology and biochemistry) from local laboratories.

All lab parameters planned in the protocol will be summarized:

- **Hematology:** absolute lymphocytes, absolute neutrophils, absolute eosinophils, absolute basophils, absolute monocytes, hemoglobin, hematocrit, WBC, RBC and platelet counts
- **Biochemistry:** AST, ALT, alkaline phosphatase, phosphate, chloride, direct and total bilirubin, creatinine, creatinine clearance, potassium (hyper and hypo), corrected calcium (hyper and hypo), fasting glucose (hyper and hypo), lipase, GGT, albumin, amylase, total cholesterol, HDL-C, LDL-C, triglycerides, magnesium (hyper & hypo), sodium (hyper & hypo), urea or blood urea nitrogen (BUN), uric acid, LDH, total protein

The following summaries will be produced for the laboratory data (by laboratory parameter):

- Number and percentage of patients with worst post-baseline CTC grade (regardless of the baseline status). Each patient will be counted only for the worst grade observed on-treatment.
- Shift tables to compare baseline to the worst post-baseline value will be produced for hematology and biochemistry laboratory parameters using:
 - CTCAE grades for laboratory parameters where CTCAE grades are defined
 - The classifications relative to the laboratory reference ranges (low/normal/high) for laboratory parameters where CTCAE grades are not defined

The laboratory assessments collected outside the on-treatment period will be flagged in the listings.

The following listings will be produced for the laboratory data:

- Listing of patients with laboratory values outside the laboratory reference ranges with values flagged to show the corresponding CTCAE grades and the classifications relative to the laboratory reference ranges
- By-patient listing of other laboratory data (collected if clinically indicated) such as urinalysis, coagulation and Serum hCG for pregnancy test .
- Listing of patients with notable laboratory abnormalities (i.e. newly occurring CTC grade 3 or 4 laboratory toxicities).

3.9.2.3 Liver function parameters

Liver function parameters of interest for BKM120 are total bilirubin (TBL), ALT, AST and alkaline phosphatase (ALP).

The following summaries will be produced:

- The number (%) of patients with worst post-baseline values in the following categories will be summarized:
 - ALT or AST > 3xULN
 - ALT or AST > 5xULN
 - ALT or AST > 10xULN
 - ALT or AST > 20xULN
 - ALP > 1.5xULN
 - TBL > 2xULN
 - ALT or AST > 3xULN & TBL > 2xULN*

ALT or AST > 3xULN & TBL > 2xULN & ALP < 2xULN*

Note: *Potential Hy's Law events (candidates) are defined as those patients with AST or ALT > 3xULN and TBL > 2xULN and ALP < 2xULN at any visit during the on-treatment period...

Note :The counts patients with ALT or AST > 3xULN, 5xULN, 10xULN etc are cumulative i.e. patient having a ALT or AST > 10x ULN would also be counted for ALT or AST > 5xULN and ALT or AST > 3xULN.Same rule applies for TBL increase as well.

A figure displaying time course of hepatic function tests (ALT, AST, TBL, ALP) in patients with Hy's law will be displayed in the safety set.

Additional categories may be added to the above list based on any updates to the internal guidelines on collection, analysis, and presentation of liver safety data.

3.9.3 Patient self-reported Mood assessments

As a safety measure for BKM120 studies, mood assessment includes two patient self-rating mood questionnaires: GAD-7 Anxiety scale and PHQ-9 Depression scale.

The primary variable for the analyses of mood will be the total scores of the two scales. Analysis will be performed based on the Safety set.

For each scale, total score will be derived by adding column scores. Categorization will be performed as described in [Table 3-5](#) to assess severity.

Table 3-5 Toxicity grading based on mood questionnaire scores

PHQ-9 (depression)		GAD-7 (anxiety)	
Score	Severity	Score	Severity
0-4	None	0-4	None
5-9	Mild	5-9	Mild
10-19	Moderate	10-14	Moderate
20-27	Severe	≥ 15	Severe

If the questionnaire is fully completed at screening and partially completed at Cycle 1 Day 1, the partially completed questionnaire will be considered as the baseline and then the total score will be missing at baseline in the analyses.

In case of missing answer to at least one item, the total score value will be missing except if available answers translate into a severe score (e.g. ≥ 20 for PHQ-9 and ≥ 20 for GAD-7). As an example: if a patient partially completes the PHQ-9 questionnaire and indicates a response of “nearly every day” (3x7) for questions 1 to 7 and leaves questions 8 and 9 unanswered, then the total score is unknown but would be ≥ 21 . As a result, it is reasonable to categorize the severity grading as “severe”.

Number (%) of patients will be described at each time point according to PHQ-9 total score categories (0-4, 5-9, 10-19, 20-27). Similar analysis will be performed to report answer to question 9 (suicide) of PHQ-9 questionnaire.

Shift tables (from Baseline to worst total score) based on the classification (0-4, 5-9, 10-19, 20-27) will be produced for PHQ-9 total score. Bar charts will be provided. Number (%) of patients will be described at each time point according to GAD-7 total score categories (0-4, 5-9, 10-14, ≥ 15). Shift tables (from Baseline to worst total score) based on the classification (0-4, 5-9, 10-14, ≥ 15) will be produced. Bar charts will be provided.

Completion status will be analyzed using a ratio of the number of completely responded assessments (i.e., no missing answer or no missing evaluation) and the number of planned assessments.

3.9.4 Vital signs

The number (%) of patients with clinically notable changes in vital signs from baseline will be presented. Clinically notable vital sign results are defined in [Table 3-6](#).

The following listing will be produced by treatment group:

- Patients with clinically notable vital sign abnormalities

Table 3-6 Clinically notable changes in vital signs

Vital Sign (unit)	Clinically notable criteria
Weight (kg)	decrease > 7% from Baseline
	increase > 7% from Baseline
Systolic blood pressure (mmHg)	≤ 90 and decrease from Baseline of >20
	≥ 180 and increase from Baseline of >20
Diastolic blood pressure (mmHg)	≤ 50 and decrease from Baseline of > 15

Vital Sign (unit)	Clinically notable criteria
Weight (kg)	decrease > 7% from Baseline
	increase > 7% from Baseline
Heart Rate (bpm)	≥ 105 and increase from Baseline of > 15
	≤ 50 and decrease from Baseline of > 15
	≥ 120 and increase from Baseline of > 15

3.9.5 Performance Status

Performance status is assessed to attempt to quantify the impact of disease on daily life activities of patients.

If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. For the baseline time window the target would be the study reference start date. If 2 assessments within a time window are equidistant from the target date (or if the closest assessment to the target date has two ECOG filled out on the same date), then the worst ECOG PS value will be used.

The performance status will be listed or summarized by time windows as defined in the study protocol and treatment arm using numbers (%) of patients in each score category. All the analyses for Performance Status will be performed on the FAS.

ECOG PS scale is used to assess physical health of patients, ranging from 0 (most active) to 5 (least active).

The performance status will be assessed according to the ECOG performance status scale (Oken 1982). ECOG performance status will be assessed at screening, at Day 1 of each cycle and at the End of Treatment assessment.

The following time-based intervals (Table 3-7) will be used to group the ECOG P.S. data over time.

Day x is defined as date of ECOG P.S. Assessment – randomization date + 1.

Table 3-7 Time based intervals to group ECOG P.S.

Visit	Time Interval
Baseline	Time period on or before the randomization date
Cycle 1	Day 2 to Day 14
Cycle 2	Day 15 to Day 42
Cycle 3,4,... Or EOT	+/- 2 weeks centered around the planned assessment i.e. Day 43, Day 70 for Cycle 3, Day 71 to Day 98 for Cycle 4 and so on... (cycle x: $(x-1)*28 - 2*7 + 1$; $(x-1)*28 + 2*7$)

Time windows are applicable for descriptive summary of ECOG P.S. data by visit. It is not applicable to time to deterioration analysis since all post-baseline assessments will be considered.

If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. For the baseline time window the target would be study day 1 (i.e., just before the patient starts randomized treatment phase); for the Cycle 1 ECOG P.S. the target would be study day 2 (C1D1+1) and thereafter it would be every 4 weeks, corresponding to the planned assessment times of day 29 (Cycle 2), day 57 (Cycle 3), etc. If 2 assessments within a time window are equidistant from the target date then the assessment obtained prior to the target date will be considered. If the closest assessment to the target date has two assessments filled out on the same date, then the worst of these assessments will be used.

Time to definitive worsening of performance status may also be analyzed.

Time to definitive deterioration of the ECOG PS is the number of days between the date of randomization and the date of the assessment at which definitive deterioration is seen. The ECOG PS deterioration is considered definitive if there is an increase in the performance status by at least one category relative to the baseline and no later improvement compared to the predefined threshold is observed within the course of the study following the time point when the deterioration was observed.

Baseline (see Section 2.1.8) is the last available assessment on or before date of randomization. If a patient has 2 ECOG PS values at the same date, the worst ECOG PS value will be taken as 'baseline'.

Example : If the ECOG PS is 1 at baseline and then 1 , 2 , 1 , 2 , 3 at D28, D57, D83, D115, and D150 respectively, then the time to definitive worsening is D115.

Example : if the ECOG PS is 1 at baseline and then 1 , 1 , 2 at D28, D57, and D83 respectively, with no assessment of the ECOG PS after D83 then the time to definitive worsening is 83 days.

If a definitive deterioration is observed after two or more missing assessments, patient is censored at the date of their last available ECOG PS assessment prior to the deterioration.

Example : If a patient has an assessment at week 6, misses the following two assessments on weeks 10 and 14 and a definitive deterioration is observed on week 15, then the event will be censored at week 6.

In addition, death is considered as a worsening of performance status if it occurs close to the last available assessment, where "close" is defined as twice the planned (i.e. protocol scheduled) period between two assessments. This avoids overestimating the time to definitive worsening in patients dying after an irregular assessment scheme. Patients who die after more than twice the planned period between two assessments are censored at the date of their last available assessment of the performance status.

Example: If the last assessment is at week 6 and the patient dies on week 10, the definitive deterioration date will be week 10. On the other hand, if the last assessment is at week 6 and the patient dies on week 16, which is after more than twice the planned period between two assessments since the last assessment (week 6), then the patient will be censored at week 6.

Patients receiving any further anti-neoplastic therapy before definitive worsening will be censored at the date of their last assessment before starting this therapy. Patients that have not worsened as of the cutoff date will be censored at the date of their last assessment before the cutoff.

Patients without baseline ECOG PS or without any post-baseline ECOG PS will be censored at the date of randomization with censoring reason being 'No baseline score' or 'No post-baseline score', respectively. However, patients without post-baseline ECOG PS* who die within D2 days after date of

randomization (which correspond to the first two planned assessments) will be counted as having a definitive deterioration of the ECOG PS at the date of death. For more details of censoring reasons for time to deterioration of ECOG PS, please refer to Section 3.8.2.4

Note: This is a protocol deviation as this is an inclusion criterion from the study protocol. The variable D2 corresponds to twice the protocol defined ECOG assessment interval plus twice the time window around each assessment. In this study, $D2 = 2 * 4 * 7 + 2 * 3 = 62$ days.

Kaplan-Meier estimates will be constructed for each treatment arm. Median for each treatment group will be obtained along with 95% confidence intervals.

ECOG PS scale is used to assess physical health of patients, ranging from 0 (most active) to 5 (least active):

Table 3-8 ECOG Performance Status

Score	Description
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair
5	Dead

A definitive deterioration will be considered for ECOG when there is an increase in the ECOG score by a margin of one unit or higher. The time to definitive deterioration of ECOG scores by treatment will be summarized using the Kaplan Meier method.

3.9.6 Cardiac assessments

The summaries will include only assessments on treatment. All assessments will be listed and those collected outside of the treatment window will be flagged.

3.9.6.1 ECG parameters

ECG data from local assessment will be summarized.

As recommended in the FDA guidance on Clinical Evaluation of QT/QTc interval prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs, the number (%) of patients having notable ECG interval values will be summarized by treatment group as follows:

Table 3-9 Clinically notable ECG values

ECG parameter (unit)	Clinically notable criteria
QT (ms), QTcF (ms)	New > 450 ms
	New > 480 ms
	New > 500 ms

ECG parameter (unit)	Clinically notable criteria
	Increase from Baseline > 30 ms
	Increase from Baseline > 60 ms
PR duration (ms)	Increase > 25% from Baseline and to PR duration > 200
QRS duration (ms)	Increase > 25% from Baseline and to QRS duration > 110
Heart Rate (bpm)	≤ 50 and decrease from Baseline of ≥ 25%
	≥ 120 and increase from Baseline of ≥ 25%

A newly occurring ECG abnormality is defined as an abnormal post-baseline ECG finding that is not present at Baseline. Baseline is defined as the last ECG measurements taken at

The percentage of patients having notable ECG interval values will be summarized by treatment and based on the total number of patients in the Safety set.

The following listing will be produced by treatment group:

- Patients with clinically notable ECG abnormalities

3.9.6.2 Cardiac imaging (MUGA / ECHO)

Shift tables comparing baseline to worst post-baseline cardiac imaging overall interpretation (MUGA or ECHO) will be provided. Percentages will be based on all patients in the Safety set. Note: If there is any change in the methodology used throughout the study compared to Baseline, the post-baseline values will be discarded in the tables presenting comparisons to Baseline.

Descriptive statistics of the left ventricular ejection fraction (LVEF) at Baseline, worst post-baseline and change from Baseline to worst post-baseline value will be provided.

The following listing will be produced by treatment group:

- Patients with newly occurring clinically significant abnormality

3.10 Subgroup analyses

3.10.1 Efficacy

PFS and OS will be summarized for each level of the stratification factor: Number of prior lines of therapy (1 vs 2) and region of investigator site (North America vs Rest of the world) . Similar analysis for PFS, OS and ORR will be performed in the subgroups of

- i. PI3K pathway status (activated vs. non-activated) based on archival tumor.



- iii. PIK3CA status (Wildtype vs Mutant) based on archival tumor and ctDNA(at screening).



For each of the subgroups, the following analyses will be performed:

- Hazard ratio with 95% C.I and Median with 95% CI of Kaplan-Meier survival estimates for PFS and OS
- ORR will be presented by treatment group along with exact 95% CIs ([Clopper & Pearson, 1934](#)).

3.11 Pharmacokinetic Analysis

FPAS will be used in the non-compartmental analysis and PAS will be used in the population PK and exposure response analysis.

The plasma samples from all patients will be assayed for BKM120 concentrations by Novartis or its designee using methods described in the Laboratory manual. Values below the assay LLOQ are assigned a value of zero in the BA source dataset. Therefore such values will be treated as zero in summary statistics and listings except for the geometric mean and geometric CV% for which they are treated as missing. Missing values will be labeled accordingly. All concentrations below the LLOQ or missing data will be labeled as such in the concentration data listings.

3.11.1 Non compartmental analysis

The aim of this PK analysis is to have a preliminary PK assessment of BKM120 exposure when administered in combination with paclitaxel in this specific population. Descriptive graphical plots of geometric mean and arithmetic mean (SD) BKM120 plasma concentration – time profiles at cycle 1 Day 15 will be generated. Additionally graphical plots of individual BKM120 plasma concentration – time profiles may be generated. Further graphical exploratory analyses may be carried out if deemed appropriate.

The following PK parameters of BKM120 :

- **AUC_{last}** : The AUC from time zero to the last measurable concentration sampling time (t_{last}) (mass x time x volume⁻¹)
- **AUC_{tau}** : The area under the plasma concentration-time curve from time zero to the end of the dosing interval, tau (mass x time x volume⁻¹)
- **C_{max}** : The maximum (peak) observed plasma drug concentration after single dose administration (mass x volume⁻¹)
- **T_{max}** : The time to reach maximum (peak) plasma drug concentration after single dose administration (time)
- **CL/F** : Apparent total body clearance of drug from the plasma after oral administration (volume x time⁻¹)
- **Rs_{qadj}** : Square of the correlation coefficient associated with λ_z

will be summarized using descriptive statistics presenting n(number of observations), arithmetic mean, SD, coefficient of variation CV (%), geometric mean, geometric CV%, median, minimum and maximum. For Tmax only median, minimum and maximum will be provided. For Rsqadj and λ_z only listings will be provided. Calculation of AUCtau will be based on the value of tau = 24 hours. Calculation of PK parameters will include up to last measurable concentration Tlast, as outlined in the Novartis Internal Guidance Standardization of pharmacokinetic parameters. Similarly summary statistics will be provided for concentration by scheduled time point including both n(number of values to be reported) and m(number of non-zero values to be reported) in addition to the summary statistics listed above.

The derived PK parameters will be presented with PK data collected in other trials with [CBEZ235A2118] or without paclitaxel [CBKM120A2101]. This analysis aims to further characterize BKM120 exposure in this specific population but no formal analysis will be conducted.

3.12 Biomarkers

The analysis of biomarker data collected at baseline will include basic descriptive statistics by treatment group and subgroup analyses with respect to key efficacy endpoints – PFS, OS and ORR for the FAS population as detailed in section 3.10.1.

For all markers with post-baseline data, baseline values/levels will be listed by patient and treatment group.

There may be circumstances when a decision is made to stop a sample collection, or not perform or discontinue the analysis of blood/archival tumor due to either practical or strategic reasons (e.g. issues related to the quality and or quantity of samples, or issues related to the assay that preclude the analysis of samples). Under such circumstances, the number of samples may be inadequate to perform a rigorous data analysis and the available data will only be listed.

3.12.1.1 Outline of Data analysis

The number/proportion (%) of patients with molecular alterations from the screening archival or fresh biopsy tumor sample will be summarized overall and by treatment group. Molecular alterations identified using Next Generation Sequencing technology is planned to be summarized for only those genes found to be altered in at least 5 % of patients, as well as key genes, such as PIK3CA, PTEN, AKT and HPV even if they do not achieve this minimum.

For analyses based on PI3K activation status (as per PIK3CA mutation, PTEN expression, see Sections 3.5.3 and 3.10.1.

Molecular alterations in ctDNA measured from the blood samples at screening will also be summarized and listed. It is planned to compare alterations identified in ctDNA at screening to those obtained in archival tumor samples.

With the collection of optional post-baseline fresh tumor biopsy samples, proteomic and genomic analysis results may be summarized. Otherwise, data may only be listed if there is sufficient sample size. Additional *post hoc* analyses of these data may occur and will be documented as part of separate biomarker report.

• **Biomarker Status**

The type of PI3K pathway activation status will be described (PIK3CA mutation, PTEN low expression). Biomarker assessment characteristics (archival tumor, fresh tumor, and their sources) will also be described.

Pi3K Activation status will be defined as following:

PIK3CA Mutation	PTEN Loss of Expression	Pi3K Activation Status
Yes (mutant)	Yes	Yes
Yes (mutant)	No	Yes
Yes (mutant)	Missing	Yes
Yes (mutant)	Unknown	Yes
No	Yes	Yes
No	No	No
No	Missing	Missing
No	Unknown	Unknown
Missing	Yes	Yes
Missing	No	Missing
Missing	Missing	Missing
Missing	Unknown	Missing
Unknown	Yes	Yes
Unknown	No	Unknown
Unknown	Missing	Missing
Unknown	Unknown	Unknown

Overall Mutation Load

Mutational load will be defined as the number of nonsynonymous alterations (mutations and short insertions/deletions) per patient. To exclude germline alterations from mutational load, alterations with an allelic fraction (AF; fraction of sequencing reads supporting the alteration) consistent with germline alterations ($0.45 \leq AF < 0.55$ or $AF \geq 0.9$) which are also of unknown functional significance will not be included. In addition, copy number alterations will be excluded in calculating mutational load.





TP53 disruptive/non-disruptive classification

TP53 alterations will be classified as disruptive versus non-disruptive based on the functional impact of the alteration on protein effect (Poeta et al, NEJM , 2007). Nonsense, frameshift and splice site acceptor/donor TP53 alterations will be classified as disruptive. In addition, nonconservative alterations (change in polarity/charge of amino acid) located inside the key DNA-binding domain (codons 163-195 or 236-251) will be classified as disruptive. All other TP53 alterations will be classified as non-disruptive. The classification of amino acids by polarity and charge can be found in the supplementary appendix of Poeta et al (NEJM, 2007). Patients will be classified into disruptive, non-disruptive and non-altered groups according to the following criteria:

TP53 alteration	TP53 category
1+ TP53 disruptive alterations	Disruptive
1+ TP53 non-disruptive alterations and 0 disruptive alterations	Non-disruptive
0 disruptive and 0 non-disruptive alterations	Non-altered

3.12.1.2 Data handling principles

PTEN loss of expression i.e PTEN Null / Low is < 10% tumor cells expressing PTEN at 1+ level (and no cell expressing PTEN at 2+ or 3+ levels).

3.13 Interim safety review analyses

A DMC will be established to monitor and review the interim safety data on an ongoing basis. The first safety review of the patients will occur 3 – 6 months after FPFV and thereafter the safety review will occur every 3 – 6 months. In addition to these periodic safety reviews the DMC will also receive and review the baseline biomarker data of patients related to their PI3K activation status. The analysis plan for interim safety and biomarker review is detailed in DMC RAP module 3.

4 General statistical methodology

4.1 Baseline comparability

Appropriate descriptive summary statistics of baseline variables (see [Section 3.5](#)) will be provided as in-text tables in the core CSR and also in the post-text tables. The summaries will be grouped by treatment group, but no p-values will be provided.

4.2 Center pooling

All study centers will be combined for the analysis unless otherwise specified. No center effect will be assessed due to expected small size of centers.

4.3 One-sided vs. two-sided test

Not Applicable.

4.4 Time-to-event analyses

The following sections present a general methodology used to analyze time-to-event variables (e.g. progression-free survival, overall survival, duration of response).

4.4.1 Analysis of time-to-event data with ties

The STRATA statement in LIFETEST procedure will be used analyze time to event data with ties. The PHREG procedure in SAS with option TIES=EXACT will be used to fit the Cox proportional hazards model.

4.4.2 Kaplan-Meier estimates

The survival function in each treatment group will be estimated using the K-M (product-limit) method as implemented in PROC LIFETEST. Median survival for each treatment group will be obtained along with 95% CIs calculated from PROC LIFETEST output. K-M estimates with 95% CIs at specific time points will be summarized. The K-M graphs as well as the statistics (median, hazard ratio etc.) displayed on the graph will be obtained from the SAS software.

General SAS code is:

```
PROC LIFETEST data=dataset;  
TIME survtime*censor(1);  
STRATA stratum1 stratum2...<stratums> /group=trt;  
RUN;  
/* stratum1 to stratums represent the stratification variables assigned at randomization (to be included for stratified analysis only);  
survtime represents variable containing event/censor times;  
censor represents censoring variable (1=censored, 0=event);  
trt represents treatment group variable; */
```

4.4.3 Hazard ratio

The HR as a measure of treatment effect will be derived from the Cox proportional hazards model using SAS procedure PHREG with TIES=EXACT option in the MODEL statement. The stratified unadjusted Cox model will be used (where the baseline hazard function is allowed to vary across strata) for the primary analysis, i.e. the MODEL statement will include only the treatment group variable as a covariate and the STRATA statement will include stratification variable(s). The strata information will be based on the data obtained from IRT that was utilized for randomization.

General SAS code for the stratified Cox model:

```
PROC PHREG data=dataset;  
MODEL survtime*censor(1)=trt / TIES=EXACT;  
STRATA stratum1 stratum2...<stratums>;  
RUN;
```

/ survtime represents variable containing event/censor times;
censor represents censoring variable (1=censored, 0=event);
trt represents treatment group variable;
stratum₁ to stratum_s represent the stratification variables assigned at randomization (to be included for stratified analysis only) */*

4.5 Bayesian methodology for the PFS and OS analyses

PFS and OS analysis will be performed in the full population using Bayesian methodology.

The PoC criteria for these two endpoints (described in [Section 3.8.1.2](#)) will be considered based on HR from a stratified unadjusted Cox proportional hazards regression model and Bayesian approaches to declare PoC for PFS and OS respectively.

Posterior distribution of logarithm of hazard ratio for PFS

Let θ denote the natural logarithm of hazard ratio (HR) in terms of the PFS of BKM120 + paclitaxel arm vs. placebo + paclitaxel arm, and $\theta < 0$ indicates efficacy in favor of BKM120 + paclitaxel.

Assume that θ follows a Normal prior distribution

$$\theta \sim N(\theta_0, 4/n_0),$$

where θ_0 is the specified prior mean and n_0 is the number of events worth of prior information.

Let y_m denote the $\log(\text{HR})$ of PFS estimated from the current study based on m observed events, then $y_m \sim N(\theta, 4/m)$.

Therefore the posterior distribution of θ is

$$\theta \sim N(\phi y_m + (1-\phi) \theta_0, 4/(m+n_0))$$

where $\phi = m/(m+n_0)$.

A non-informative prior (HR=1 and $n_0=0$) is used to derive the posterior distribution of θ given above; i.e.,

$$\theta \sim N(y_m, 4/m)$$

Posterior distribution of logarithm of hazard ratio for OS

The posterior distribution for the logarithm of hazard ratio for OS will be estimated in the same way described above for PFS using the overall survival data.

4.6 Confidence intervals for response rate and disease control rate

Response rate and disease control rate will be summarized as percentages with 95% confidence intervals. As a standard, an *exact binomial confidence interval* (implemented using SAS procedure FREQ with EXACT statement for one-way tables) will be calculated ([Clopper & Pearson, 1934](#)).

5 References

Brookmeyer R, Crowley J (1982) A confidence interval for the median survival time. *Biometrics*, 38: 29-41.

Clopper, C.; Pearson, E.S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*; 26, 404-413.

Collet D (1994). *Modelling survival data in medical research*. London, Chapman & Hall.

CPMP/EWP/2863/99 (2003). Points to Consider on Adjustment for Baseline Covariates

FDA (2007). *Guidance for Industry Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics*, US Department of Health and Human Services.

FDA (2005), *Guidance for industry E4 Clinical Evaluation of QT/QTc interval prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs*,

International Conference on Harmonization. E9: Statistical Principles for Clinical Trials.

Schemper, M.; Smith, TL (1996). A Note on Quantifying Follow-up in Studies of Failure Time. *Controlled Clinical Trials*; 17:343-346

M. Luana Poeta, M.D, et al. (2007) *TP53* Mutations and Survival in Squamous-Cell Carcinoma of the Head and Neck. *N Engl J Med*. 2007 December 20; 357(25): 2552–2561.