NOVARTIS

Clinical Development

BYL719/alpelisib

CBYL719F12201/ NCT04589650

EPIK-P2: A Phase II double-blind study with an upfront, 16-week randomized, placebo-controlled period, to assess the efficacy, safety and pharmacokinetics of alpelisib (BYL719) in pediatric and adult patients with PIK3CA-related overgrowth spectrum (PROS)

Statistical Analysis Plan (SAP)

Document type:	SAP Documentation
Document status:	Final Amendment 3
Release date:	29-Apr-2024
Number of pages:	70

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
25-Mar-2021	Prior to FPFV	Creation of final version	N/A - First version	NA
10-Oct-2022	Prior to DB lock	Creation of amendment 1	Changes made to align with study Global Protocol Amendments 1 (PA1) and 2 (PA2), which introduced the following main changes: 1) Update of design elements, specifically the hypothesis to be tested for the primary endpoint, the sample size and the addition of a secondary objective to assess the time to treatment failure [PA1] 2) Modification of the primary objective and related endpoint to confirmed objective response [PA2] 3) Inclusion of an additional analysis at Month 46 to characterize the durability of confirmed objective response [PA02] 4) Add an exploratory group (Group 4) of approx. 6 participants, 2-5 years of age, treated with alpelisib film-coated tablets (FCT) [PA1] 5) Add an exploratory group (Group 5) of approx. 15 participants, 6-17 years of age, treated with a starting dose of 125mg alpelisib FCT [PA2] 6) Modify the imaging assessment frequency [PA1, PA2]	To align with study Global Protocol Amendments 1 and 2, the following changes were made: Section 1: clarified the timing of the Week 48 CSR and added an additional CSR at Month 42. Section 1.1: updated design as per PA2 Section 1.2: updated objectives and primary estimands as per PA2 Section 2.1: Added definition of analysis period for the exploratory groups. Clarified treatment discontinuations for the analysis period stop date in Table 2-1. Section 2.1.1: added general definitions for the exploratory groups. Example added to further clarify windows for multiple assessments. Section 2.2 and 2.2.1: updated analysis sets and classification rules for exploratory groups Section 2.3: added details on demographics and disposition for exploratory groups Section 2.4: added details on study treatment for exploratory groups Sections 2.5, 2.5.1: updated primary objective and related endpoint to confirmed response as per PA2. Section 2.5.2: updated primary hypothesis (and sample size) as per PA1. Sections 2.5.3, 2.5.4: updated handling of intercurrent events and missing values as per PA2. Section 2.6.1: updated key secondary efficacy variable as per PA2. Section 2.7.2: clarified that analyses are done by age group. Section 2.7.5: clarified that duration of response will be analyzed for participants with a confirmed response. Section 2.7.6: added time to treatment failure analysis as per PA1.

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
				Section 12.11: clarified applicable age-range of the PRO measures for analysis as per PA1.
				[redacted]
				Section 3: updated sample size and operating characteristics as per PA1 and PA2.
				Section 5.5.2.4: clarified applicable age-range of the PRO measures as per PA1. Added missed cognitive function domain.
			Other changes (not related to PA1 or PA2)	Other changes made: Section 2.1: clarified that analysis details for Group 3 will be specified in a separate SAP. Section 2.1.1: clarified that all analyses will be presented by age group, unless otherwise stated. Section 2.5.1: added reference regarding the handling of multiple imaging dates for a specific visit. Section 2.5.6: clarified that other sensitivity analyses may be done as needed. Section 2.6.1: clarified the method to be used for the computation of the confidence interval for the difference in proportions. Section 2.8.1: added reference on search criteria for SARS-CoV-2 infections. Section 2.8.1.2: clarified that time to and duration of first occurrence of AE will be done only if there is a sufficient number of events. Removed analysis of time to first occurrence of diarrhea (PT) and rash (PT), as a respective analysis of GI toxicity (AESI) rash (AESI) is already specified. Accordingly, removed the duration of first occurrence of ≥Grade 2 diarrhea (PT). Section 2.11: removed additional analyses of clinically important change in dyspnea items, as this is a secondary PRO endpoint and not of primary interest. Section 2.13.8: removed details for Group 3, as not applicable for this primary analysis SAP. Section 5.1.3 added rules for handling of multiple imaging dates for PROS lesion assessments.

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
				Section 5.4.1: added age-specific notable ECG values for participants <18 years.
			Edits made to correct typo or improve clarity/ consistency of the document.	
28-Jul-2023	Prior to DB lock	Creation of amendment 2	Changes made to align with study Global Protocol Amendment 3, which introduced the following main changes (no impact on analysis): 1) Add a new exclusion criterion 2) Add physical exam on Day 1	Section 1: updated reference to Global Protocol Amendment 3.
			Updates made to improve clarity	Section 2.2.3: clarified that the term “PROS syndromes” denotes “PROS phenotypes”. Added reference to separate SAP for Chinese subgroup analysis. Section 2.4.1: clarified the definition of dose interruptions. Section 2.7.1: updated generation of waterfall plots as appropriate. Section 2.7.2: clarified that analyses of change from baseline in participants with impairment at baseline will be done as applicable. Section 2.8.1: clarified the definition of treatment-emergent events. Separated summaries of AEs leading to dose reduction and to dose interruption.
			Updates made to align with Novartis statistical guidance on growth and development	Section 2.8.4.3: re-structured the section for clarity/alignment with the Novartis guidance regarding the analysis of sexual maturation; removal of analyses of Marshall standards Section 6: references updated with regards to updates of Section 2.8.4.3.
29-Apr-2024	Prior to DB lock	Creation of amendment 3	Changes made to align with study Global Protocol Amendment 4 (PA4), which introduced the following main changes: 1) Modified the timing of the primary analyses for Groups 1 and 2 to occur after all participants of those groups have completed at least 48	Section 1: updated reference to Global Protocol Amendment 4. Section 2.5.6: added supportive analysis of primary endpoint according to initial cut-off (up to 24 weeks).

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			weeks (instead of 24 weeks) Other changes (not related to PA4)	Other changes made: [redacted] Section 2.1: added note that some analyses will present data during overall study period. Clarified the derivation of the end of the analysis periods in Table 2-1 and Section 2.1.1.8. Section 2.1.1.9: added further clarification for the time window derivation. Section 2.3.1: removed analysis of PDs by analysis period. Section 2.4.1: added definition of dose reductions, increases and overdoses. Section 2.7.1: added clarification about derivation of response at assessment visits. Section 2.7.2: added analysis of shift in CTCAE grades of PROS-related medical conditions over time. Section 2.7.5: added clarifications for the start date for the duration of response. Section 2.8.1.2: removed analysis by placebo-controlled period. Further clarification that a sufficient number of events is needed for the analysis. Section 2.9: added PK parameters (AUC0-24h, AUC%Extrap, Rsq_adj) for evaluation if data permit. Section 2.11: clarified analysis of weekly average PROs from the diary (Wong-Baker Faces, Dyspnea items). Section 5.4.1: clarified the baseline definition of age-specific ECG HR notable values to align with the definition in EPIK-P3 study. Section 5.4.3: updated the classification of pain medications to be based on ATC level 4 codes to align with the coding in the database. Clarified that categorization into weak and strong opioids will be done in case of sufficient data. [redacted]

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
				Section 5.5.3: details added about further time intervals for the weekly averages.
			Edits made to correct typo or improve clarity/ consistency of the document.	Section 6: references removed with regards to updates in Section 5.5.2.

Table of contents

List of abbreviations

2MWT	2 minute walk test
AE	Adverse event
ATC	Anatomical Therapeutic Classification
AUC	Area Under the Curve
BIRC	Blinded independent review committee
BPI	Brief Pain Inventory
CI	Confidence interval
CRF	Case Report Form
CSR	Clinical Study report
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
DMC	Data Monitoring Committee
eCRF	Electronic Case Report Form
eCRS	electronic Case Retrieval Strategy
FAS	Full Analysis Set
[redacted]	
GFR	Glomerular Filtration Rate
IVR	Interactive Voice Response
LVEF	Left Ventricular Ejection Fraction
MDRD	Modification of Diet in Renal Disease
MedDRA	Medical Dictionary for Drug Regulatory Affairs
NCI	National Cancer Institute
p.o.	Per os / orally
PAS	Pharmacokinetic analysis set
PK	Pharmacokinetics
PRO	Patient-reported Outcomes
PROS	PIK3CA-related overgrowth spectrum
PROMIS	Patient Reported Outcome Measurement Information System
QoL	Quality of Life
RAP	Reporting & Analysis Process
SAP	Statistical Analysis Plan
SOC	System Organ Class
TFLs	Tables, Figures, Listings
WHO	World Health Organization

1 Introduction

This statistical analysis plan (SAP) describes the planned analyses for the primary Clinical Study Report (CSR) of the study CBYL719F12201, a Phase II multi-center study with an upfront 16-week, randomized, double-blind, placebo-controlled period, and extension periods, to assess the efficacy, safety and pharmacokinetics of alpelisib in pediatric and adult participants with PROS.

The primary analysis will be performed after all participants (Groups 1 and 2) have completed at least 48 weeks of study treatment or have discontinued earlier. Results will be presented in the primary CSR.

An additional CSR will be developed once all participants in Groups 1, 2, 4 and 5 have reached at least 48 months (i.e., 36 months after the Week 48 timepoint) of treatment or discontinued earlier to characterize the durability of response. Furthermore, additional reports may also be developed prior to the final clinical study report in order to respond to Health Authority requests or to support Health Authority submissions. Any additional data for participants continuing to receive study treatment past the data cutoff date for the primary CSR/interim CSRs, as allowed by the protocol, will be reported at completion of the study in a final CSR. The corresponding analyses for those CSR(s) will be described in separate SAP(s).

The content of this SAP is based on the CBYL719F12201 protocol amendment v04 released on Oct 5th, 2023. All decisions regarding the analysis, as defined in the SAP document, have been made prior to database lock and unblinding of the study data.

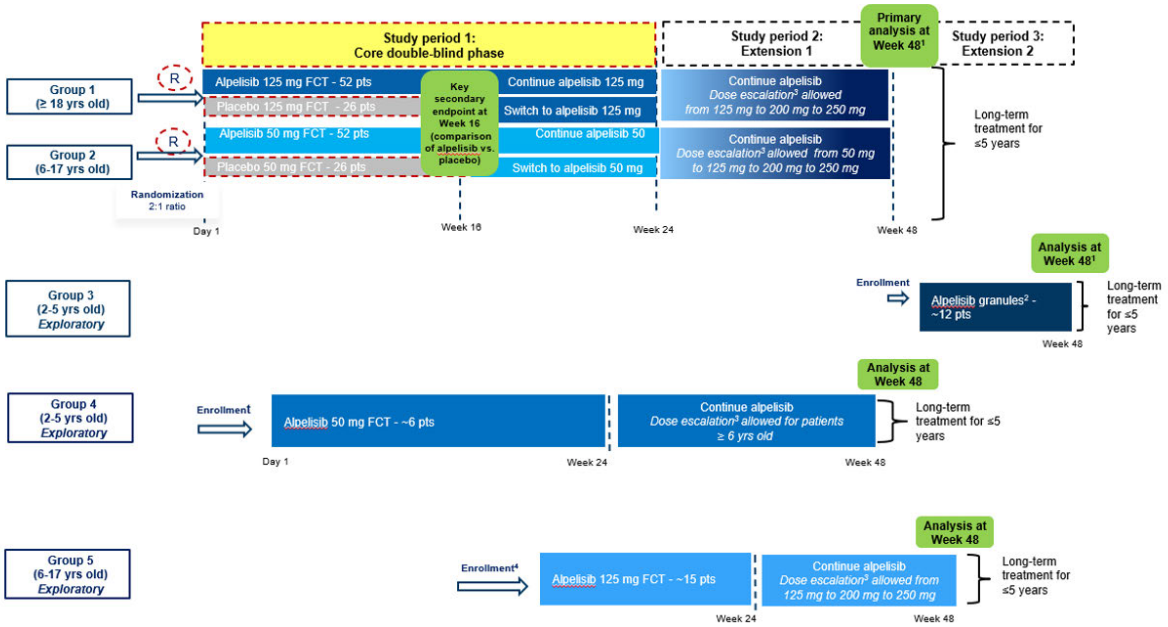
1.1 Study design

The study is a Phase II multi-center study with an upfront 16-week, randomized, double-blind, placebo-controlled period, and extension periods, to assess the efficacy, safety and PK of alpelisib in pediatric and adult participants with PROS.

A total of approximately 156 participants (of age ≥ 6 years) will be randomized in the study from Groups 1 and 2. Three additional groups with a total of approximately 33 participants will be enrolled for exploratory purposes (see Figure 1-1):

- Group 3 (2 to 5 years old) of approximately 12 participants will receive the alpelisib granules formulation in an open-label setting. The dose of the granules will be determined based on the primary analysis for efficacy, safety and PK of alpelisib in Groups 1 and 2 in addition to the data from Groups 4 and 5 as available.
- Group 4 (2 to 5 years old) of approximately 6 participants will receive a starting dose of 50 mg alpelisib FCT formulation once daily in an open-label setting.
- Group 5 (6 to 17 years old) of approximately 15 participants will receive a starting dose of 125 mg alpelisib FCT formulation once daily in an open-label setting.

Figure 1-1 Study design



[1] Study level milestones once all participants have completed the specified time point or discontinued earlier

[2] Starting dose for Group 3 will be defined based on the analysis of available data at the time of the primary analysis of the primary endpoints (i.e. when all participants from Groups 1 and 2 have completed 48 weeks of treatment or discontinued earlier).

[3] Refer to protocol Section 6.5 for details on alpelisib dose escalation and dose modification.

[4] Enrollment in Group 5 will start after implementation of Global Protocol Amendment 02 and enrollment into Group 2 has been completed

Core, Extension 1 and Extension 2 periods will be applied to the participants ≥6 years old (Group 1 – adults, Group 2 – participants 6 to 17 years old).

Group 3 and 4 will include the participants who are 2 to 5 years old and will be exploratory groups.

Group 5 will include participants who are 6 to 17 years old and will be an exploratory group.

Group 1 and Group 2 will be enrolled in parallel.

Group 3 will be open to enrollment after the primary analysis of participants in Groups 1 and Group 2 will be available in addition to the data from Groups 4 and 5 as available and only after implementation of a future substantial Global Protocol Amendment.

Group 4 will be open to enrollment immediately after the implementation of Global Protocol Amendment 01.

Group 5 will be open to enrollment after implementation of Global Protocol Amendment 02 and immediately after enrollment of Group 2 has been completed.

Study period 1 - Core Period: Double-blind treatment, with an upfront 16-week placebo-controlled period (From Randomization to the end of Week 24) – Groups 1 and 2

At study start the participants of Group 1 and Group 2 will be enrolled. Eligible participants aged ≥6 years old will be randomized in a 2:1 ratio to alpelisib or matching placebo; the starting dose will be 125 mg once daily p.o. in Group 1, and 50 mg one daily p.o. in Group 2. The upfront placebo-controlled period will continue for the first 16 weeks. At the conclusion of Week 16 those participants who were randomized to receive placebo will be switched to active treatment with alpelisib and will receive 125 mg once daily p.o (Group 1) and 50 mg once daily p.o (Group 2) in a blinded fashion. Those participants who were randomized to receive alpelisib, will continue their treatment at the same dose level. No information regarding the treatment received during the first 16 weeks will be shared at the time of the switch to ensure maintaining the blinding.

Dose reductions because of safety/tolerability issues will be allowed in all groups at any time.

Dose escalation is not allowed for Group 1 and Group 2:

- During Core period (Week 1 – Week 24).
- First 4 weeks of Extension 1 period (Week 25 – Week 28)

See protocol Section 6.5.1.

The randomized treatment assignment will remain blinded to participants, investigators and the study team until the time of the primary analysis, when the last participant completes week 48 or discontinues earlier.

Study period 2 - Extension 1: treatment with alpelisib (Week 25 up to the end of Week 48) – Groups 1 and 2

Participants (Group 1 and Group 2) will continue their treatment at the assigned dose if this dose was not reduced due to safety-tolerability reasons. Participants, who after completion of the first 24 weeks of study treatment are taking a dose lower than the initial one at randomization due to safety-tolerability issues, should continue on the same (reduced) dose. Once a participant has reached week 29 then dose escalation will be allowed as per protocol Section 6.5.1.

Dose escalation is NOT allowed during first 4 weeks of Extension 1 period (Weeks 25-28) to ensure that participants who were randomized to placebo have received at least 12 weeks of alpelisib at initially assigned dose when dose escalation can be considered.

Study period 3 - Extension 2: long-term treatment with alpelisib (Week 49 up to 5 years) – Groups 1 and 2

Groups 1 and 2 participants who continue the study treatment until Week 48 and continue deriving clinical benefit as determined by the investigator, will enter a long-term extension period. Additional safety and efficacy data will be collected during the extension period. This period will last up to 5 years from randomization.

Exploratory study part:

- **Group 3, open-label treatment with the alpelisib granules formulation:**

Group 3 participants will receive alpelisib in an open-label setting. The granules formulation currently under development will be used in this group. The dose of the granules will be determined based on the primary analysis for efficacy, safety and PK of alpelisib in Groups 1 and 2 in addition to the data from Groups 4 and 5 as available. An extrapolation approach will

be used for dose selection for this group. Novartis will amend the study protocol to provide instruction on the starting dose and dose modification rules for this group in the future, in the context of a substantial Global Protocol Amendment.

- **Group 4, open-label treatment with the alpelisib FCT formulation:**

Participants of Group 4 will be enrolled immediately after implementation of Global Protocol Amendment 01 and will receive 50 mg alpelisib FCT formulation once daily in an open-label setting.

Dose escalation is allowed for those who did not derive sufficient clinical benefit at the investigator's discretion, once the participant has reached the age of 6 years old and has reached at least Week 25. Further dose escalation is allowed at a minimum of 12-week intervals until the end of the study, according to the dose modification guidance. Dose reductions because of safety/tolerability issues will be allowed at any time. Refer to protocol Section 6.5 for further details.

- **Group 5, open-label treatment with the alpelisib FCT formulation:**

Participants of Group 5 will be enrolled immediately after implementation of Global Protocol Amendment 02 and enrollment of Group 2 has been completed and will receive a starting dose of 125 mg alpelisib FCT formulation once daily in an open-label setting.

Dose escalation is allowed for those who did not derive sufficient clinical benefit at the investigator's discretion, once the participant has reached at least Week 25. Further dose escalation is allowed at a minimum of 12-week intervals until the end of the study, according to the dose modification guidance. Dose reductions because of safety/tolerability issues will be allowed at any time. Refer to protocol Section 6.5 for further details.

End of Study for all Groups:

End of Study (EoS) will occur when all participants have completed 5 years of treatment, unless the participant discontinues earlier.

Data Monitoring Committee (DMC)

A DMC will review safety data at regular intervals. For any details, please refer to the DMC Charter as well as the separate planning documents for the DMC analyses. The DMC analyses will use this SAP as basis with regard to rules, definition of endpoints, and statistical analyses.

1.2 Study objectives and endpoints

Table 1-1 outlines the primary, secondary [redacted]objectives and related endpoints. Primary, secondary [redacted] will be applicable for participants in Group 1 and 2; [redacted]

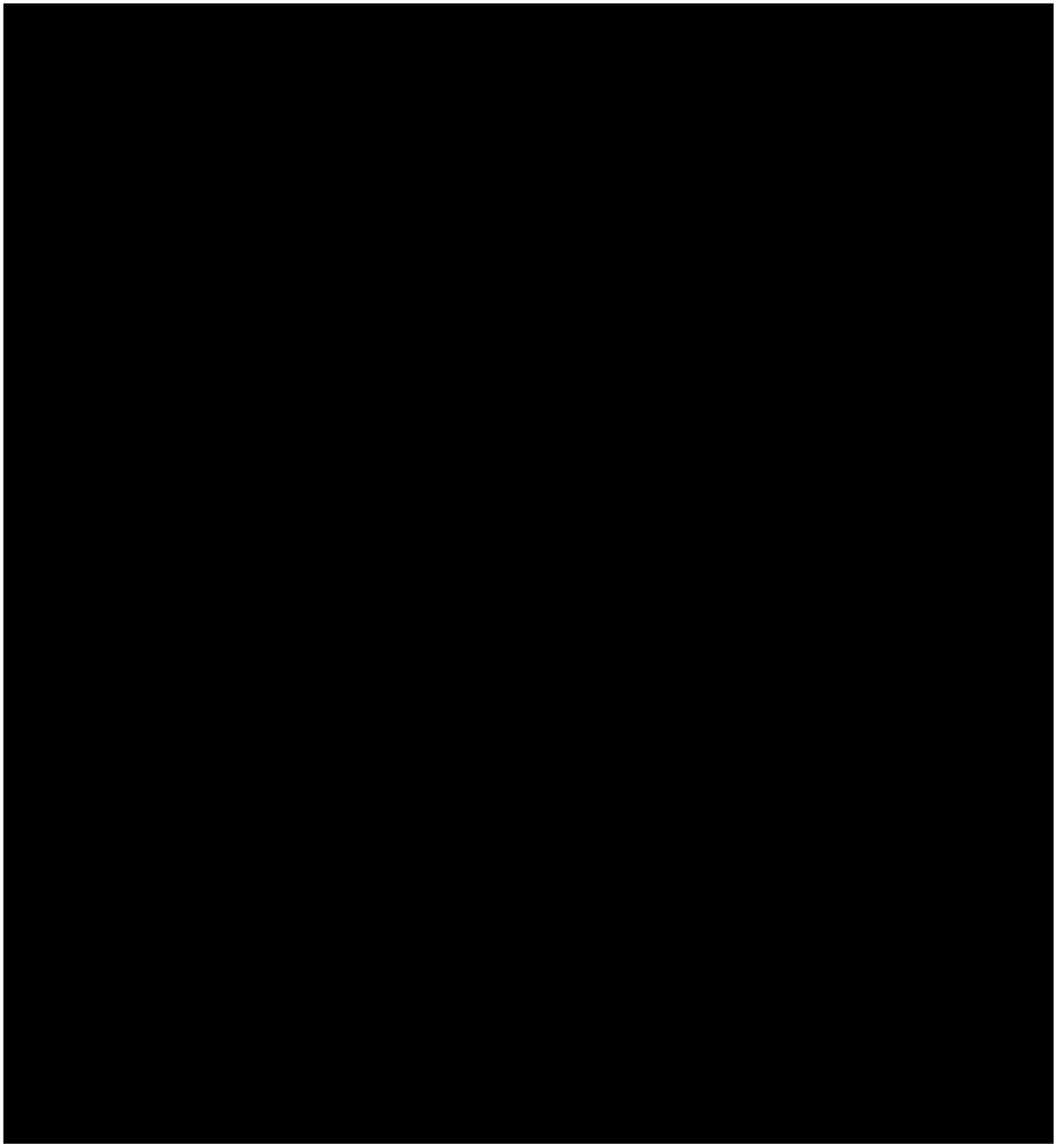
Table 1-1 Objectives and related endpoints

Objective(s)	Endpoint(s)
Primary objective(s)	**Endpoint(s) for primary objective(s)**
• To demonstrate the efficacy of alpelisib as measured by the	• Response (yes/no) defined by achieving at least 20% reduction from baseline in the sum of target

Objective(s)	Endpoint(s)
proportion of participants randomized to alpelisib with a confirmed objective response (by BIRC) in at least one of the following groups: • Group 1 (≥ 18 yr-old) • Group 2 (6 - 17 yr-old)	lesion volumes (1 to 3 lesions, assessed by MRI by a blinded independent review committee (BIRC)), provided that none of the individual target lesions has ≥ 20% increase from baseline and in absence of progression of non-target lesions and without new lesions. Confirmation of response requires a subsequent imaging assessment performed at least 4 weeks after the onset of response. See Section 1.2.1 for Primary Estimands and protocol Section 8.3 for response definition.
Secondary objective(s)	**Endpoint(s) for secondary objective(s)**
• **Key secondary objective** To demonstrate the efficacy of alpelisib vs placebo based on the comparison of the proportion of participants with response at Week 16 in Group 1 or Group 2	• **Endpoint for key secondary objective** Response at Week 16 (by BIRC): See Section 1.2.2 for secondary estimands and protocol section 8.3 for response definition.
• **Other secondary objectives**	• **Endpoints for other secondary objectives**
• To assess the efficacy of alpelisib as measured by the proportion of participants with response at Week 24 (by BIRC) in Groups 1 and 2	• Response at Week 24 (by BIRC). See protocol section 8.3 for response definition.
• To assess safety and tolerability of alpelisib as compared to placebo in Groups 1 and 2 up to Week 16	• Incidence, type, and severity of treatment-emergent adverse events per CTCAE v4.03 criteria and other safety data including changes in laboratory values, vital signs, assessments of cardiac and lung functions
• To assess the overall safety and tolerability of alpelisib in participants with PROS over time	• Incidence, type and severity of treatment-emergent adverse events per CTCAE v4.03 criteria and other safety data including changes in laboratory values, vital signs, assessments of cardiac and lung functions, growth, bone/dental development and sexual maturation (for applicable age)
• To assess changes in patient-reported pain intensity and overall severity of symptoms at Week 16 on treatment with alpelisib as compared to placebo in pediatric and adult populations.	• Change from baseline to Week 16 in Brief Pain Inventory (BPI) Worst Pain Intensity Item of the PRO (Patient reported Outcomes) diary and Patient Global Impression of Symptom Severity in pediatric (12 to 17 yr old) and adult populations (≥ 18 yr-old). See Protocol section 8.5.1 for more details on Clinical Outcome Assessments.

• To assess changes in target and non-target lesions over time and appearance of new lesions on treatment from baseline over time	• Change from baseline (as assessed by BIRC) in the:
	  • Sum of target lesion volume
	  • Sum of MRI-measurable non-target lesion volume
	  • Sum of all MRI-measurable (target and non-target) lesion volume
	• Change in other non-target lesions (by BIRC)
	• Appearance of new lesions (by BIRC)

Objective(s)	Endpoint(s)
• To assess the PK of alpelisib in adult and pediatric patients with PROS	• PK parameters (e.g. Cmax, Ctrough etc.) in Groups 1 and 2
• To assess changes in patient-reported pain, health-related quality of life and overall impression of symptoms in pediatric and adult populations over time	• Change in scores from BPI items, or Wong-Baker Faces Scale (age appropriate), PROMIS-profile (Patient Reported Outcome Measurement Information System) and Patient Global Impression of Symptom Severity
• To assess the duration of response in participants who receive alpelisib	• Duration of response (see Section 2.7.5) in participants who receive alpelisib and who satisfy the response criteria (by BIRC).
• To assess the time to treatment failure in participants who are on treatment with alpelisib	• Time to treatment failure (see Section 2.7.6) in participants who are on treatment with alpelisib
• To assess the rate of overall clinical response as assessed by Investigator at the scheduled protocol visits for disease evaluation (e.g. Week 16, 24, 40, 48, 72, 96) and thereafter every 48 weeks	• Overall clinical response as assessed by Investigator (see protocol Section 8.3 for definition)
• To assess the proportion of participants with a response at the scheduled protocol visits for disease evaluation during the extension periods	• Response (yes/no) at scheduled protocol visit. See protocol Section 8.3 for response definition
• To assess changes in symptoms and complications/comorbidities up to Week 16 on treatment with alpelisib as compared to placebo.	• Change in PROS-related symptoms and complications/comorbidities up to Week 16 among participants with symptoms and complications/comorbidities present at baseline
• To assess changes in symptoms and complications/comorbidities associated with PROS over time.	• Change in PROS-related symptoms and complications/comorbidities among participants with symptoms and complications/comorbidities present at baseline
• To assess the frequency of healthcare visits/hospitalizations due to PROS, rescue surgeries for PROS (incl. avoidance/delay in planned disease-related surgery) over time	• Number/percentage of participants with healthcare visits/hospitalized due to PROS; number of hospitalizations • Number/percentage of participants with surgeries required to manage PROS; number of surgeries



1.2.1 Primary estimands

The estimand is the precise description of the treatment effect and reflects strategies to address events occurring during trial conduct which could impact the interpretation of the trial results (e.g., premature discontinuation of treatment).

The primary scientific question of interest is, for children/adolescents aged 6 to 17 years (in Group 2) and adults (Group 1: ≥ 18 years) with PROS, to assess the benefit of alpelisib with regards to the proportion of confirmed responders (see protocol section 8.3 for definition) by BIRC, considering participants that discontinue treatment prior to confirmation of response and

participants that receive surgery as rescue therapy for any PROS lesions prior to confirmation of response as non-responders.

The primary estimands are characterized by the following 5 attributes:

1) Treatment: the randomized treatment with daily alpelisib, regardless of dose changes, in absence of rescue surgery for PROS lesions (see protocol section 6.2).

2) Population: Children/adolescents aged 6 to 17 years and adults aged ≥ 18 years with PROS (as per the inclusion/exclusion criteria)

3) Intercurrent events

- participants discontinuing treatment prior to confirmation of response: Classified as a non-responder
- participants receiving surgery as rescue therapy for any PROS lesions (target or other) prior to confirmation of response: Classified as a non-responder

4) Variable: response (yes/no) defined by achieving ≥ 20% reduction from baseline in the sum of target lesion volumes (1 to 3 lesions, assessed by MRI by BIRC), provided that none of the individual target lesions has ≥ 20% increase from baseline and in absence of progression of non-target lesions and without new lesions. Confirmation of response requires a subsequent imaging assessment performed at least 4 weeks after the onset of the response.

5) Summary Measure: Proportion of participants (children/adolescents and adults) achieving confirmed objective response (by BIRC).

1.2.2 Secondary estimands

The key secondary scientific question of interest is, for children/adolescents aged 6 to 17 years (in Group 2) and adults (Group 1: ≥ 18 years) with PROS to assess the benefit of alpelisib as compared to placebo with regards to the proportion of responders (see protocol section 8.3 for definition) at Week 16, considering participants that discontinue treatment prior to Week 16 and participants that receive surgery as rescue therapy for any PROS lesions as non-responders.

The key secondary estimands are characterized by the following 5 attributes:

1) Treatment: the randomized treatment (daily alpelisib or placebo), regardless of dose changes, in absence of rescue surgery for PROS lesions (see protocol section 6.2)

2) Population: Children/adolescents aged 6 to 17 years and adults aged ≥ 18 years with PROS (as per the inclusion/exclusion criteria)

3) Intercurrent events

- participants discontinuing treatment prior to Week 16: Classified as a non-responder
- participants receiving surgery as rescue therapy for any PROS lesions (target or other) prior to Week 16: Classified as a non-responder

4) Variable: response (yes/no) defined by achieving ≥ 20% reduction from baseline in the sum of target lesion volumes (1 to 3 lesions, by BIRC) at Week 16, provided that none of the individual target lesions has ≥ 20% increase from baseline at Week 16 and in absence of progression of non-target lesions and without new lesions

5) Summary Measure: Difference in proportion of participants (children/adolescents and adults) achieving response at Week 16.

2 Statistical methods

2.1 Data analysis general information

The primary analysis specified in this SAP will be performed by Novartis. SAS version 9.4 or later and R version 3.4.3 or later will be used to perform all data analyses and to generate tables, figures and listings.

Pooling of centers: Unless specified otherwise, data from all study centers will be pooled for the analysis. Due to expected small number of participants enrolled at centers, no center effect will be assessed.

Qualitative data (e.g., gender, race) will be summarized by means of contingency tables; a missing category will be included as applicable. Percentages will be calculated using the number of participants in the relevant population or subgroup as the denominator.

Quantitative data (e.g., age, body weight) will be summarized by appropriate descriptive statistics (e.g. mean, standard deviation, median, 25th and 75th percentiles, minimum, and maximum).

For the primary analysis a data cut-off date will be defined after all participants (group 1 and 2) have completed 48 weeks of study treatment or have discontinued earlier. All statistical analyses will be performed using all data collected in the database up to the data cut-off date. All data with an assessment date or event start date (e.g. laboratory assessment date or start date of an adverse event) prior to or on the cut-off date will be included in the analysis. Any data collected beyond the cut-off date will not be included in the analysis and will not be used for any derivations.

The analysis for the primary CSR (including the data from core period defined in Section 1.1) will comprise the following analysis periods:

Groups 1 and 2:

1. Placebo-controlled period: includes the upfront placebo-controlled 16 week period where participants will be either randomized to alpelisib or placebo. Assessment and events from randomization/baseline and up to Week 16 will be included in the analyses presented for the analysis period.
2. Alpelisib period: includes data collected on participants treated with alpelisib, i.e., participants randomized to alpelisib and participants that switch to alpelisib after treatment with placebo. Assessments and events while on treatment with alpelisib and up to the data cut-off date for the primary analysis will be included in the analysis period.

Some analyses will present data for the overall study period (from randomization/treatment start date up to cut-off date).

[redacted]

Groups 4 and 5:

1. Alpelisib period: includes data collected on participants assigned to/treated with open-label alpelisib. Assessments and events from treatment start up to data cut-off date will be included in this analysis period.

Group 3 will be open to enrollment after the primary analysis of participants in Group 1 and Group 2. Details for the analysis of this group will be specified in a separate CSR SAP. For the remainder of this SAP, details for Group 3 will not be included.

For Group 5, an analysis period to include the first 16 weeks of treatment will be defined as per the analysis conventions for Group 2.

In general, data with an assessment date or event start date (e.g., vital sign assessment date or start date of an adverse event) within an analysis period start and stop dates will be summarized for that period. The start and stop dates for the analysis periods are defined in Table 2-1.

At the time of study initiation/enrollment, the COVID-19 pandemic is ongoing. Further analysis periods (e.g., during pandemic, after pandemic) may be defined to assess the impact of the pandemic on the study results as applicable.

Table 2-1 Analysis periods (for efficacy and safety analyses)

Placebo-controlled period			Alpelisib period		
Treatment	Start date*	Stop date+	Treatment	Start date	Stop date+
Alpelisib	Randomization date or start date of alpelisib	Minimum[Date of treatment discontinuation, Week 17 Day 1 visit date – 1 day, date of death]	Alpelisib (initially randomized to alpelisib)	Randomization date or start date of alpelisib	Minimum[Date of treatment discontinuation, Cut-off date, date of death]
Placebo	Randomization date or start date of placebo	Minimum[Date of treatment discontinuation, Date of first administration of alpelisib – 1 day, date of death]	Alpelisib (initially randomized to placebo)	Date of first administration of alpelisib	Minimum[Date of treatment discontinuation, Cut-off date, date of death]

* Start date for analyses based on the Full Analysis Set is relative to Randomization date, start date for analyses based on Safety set, Pharmacokinetic Analysis Set is relative to start of study treatment.

+ For safety analyses, data up to the date of treatment discontinuation + 30 days will be considered for participants who discontinue treatment prior to the stop date of the analysis period

Note: the first non-zero dose of open-label alpelisib is considered as Week 17 Day 1 visit date in the above derivation for randomized participants in Groups 1 and 2.

Note: MRIs from End of Treatment visit for discontinued participants will be considered in the efficacy analysis for derivation of e.g. duration of response.

2.1.1 General definitions

Unless otherwise stated, all analyses will be presented by age group (i.e., Groups 1, 2, 4, 5). Similarly, the conventions and analysis rules for Group 2 also apply to the exploratory Groups 4 and 5.

2.1.1.1 Study treatment and investigational drug

In this study, the "***study treatment***" refers to alpelisib or placebo (received during any study period).

The term ***"investigational drug"*** refers to the Novartis study drug, alpelisib/BYL719.

2.1.1.2 Date of first administration of investigational drug

The date of first administration of investigational drug is defined as the first date when a non-zero dose of investigational drug is administered and recorded on the study treatment (e)CRF page.

For participant randomized to placebo, the date of first administration of investigational drug will be the first non-zero dose of alpelisib after the upfront 16-week placebo-controlled period.

2.1.1.3 Date of last administration of investigational drug

The date of last administration of the investigational drug is defined as the last date when a non-zero dose of the investigational drug is administered and recorded on the study treatment (e)CRF page. This date will also be referred as *last date of investigational drug*.

2.1.1.4 Date of first administration of study treatment

The ***date of first administration of study treatment*** is defined as the first date when a non-zero dose of study treatment (alpelisib or placebo) is administered.

Furthermore, the date of first administration of study treatment will also be referred as ***start date of study treatment.***

For participants randomized to placebo, the start date of placebo treatment is defined as the first date when a non-zero dose of placebo is administered. For participants switching to alpelisib after week 16, the start date of alpelisib treatment is defined as the first date when a non-zero dose of alpelisib is administered.

2.1.1.5 Date of last administration of study treatment

The ***date of last administration of study treatment*** is defined as the last date when the last non-zero dose of the study treatment (alpelisib or placebo) is administered.

For participants randomized to placebo, the ***date of last administration of placebo treatment*** will be the last non-zero dose of placebo during the 16 week placebo-controlled period. For participants switching to alpelisib after week 16, the **date of last administration of alpelisib treatment** will be the last non-zero dose of alpelisib.

2.1.1.6 Study day

The study day describes the day of the event or assessment date, relative to the reference start date (randomization date, start date of study treatment, start date of the investigational drug).

The study day is calculated as follows:

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

Groups 1 and 2: In the overall study period and in the placebo-controlled period the reference start date for safety assessments (e.g., adverse event onset, laboratory abnormality occurrence,

vital sign measurement, dose adjustments, PK etc.) is the start of study treatment (alpelisib or placebo). In the Alpelisib period the reference start date for safety assessment is the start of the investigational drug (alpelisib).

Groups 1 and 2: The reference start date for all other non-safety assessments (e.g., PROS lesion assessment, [redacted], Karnofsky/Lansky performance status, patient reported outcomes (PRO)) is the date of randomization in the overall study period and in the placebo-controlled period. In the Alpelisib period the reference start date for these assessment is the start of the investigational drug for participants that switch to alpelisib.

Groups 4 and 5: The reference date for all assessments (safety, efficacy, PK, PROs, etc.) is the start of study treatment (alpelisib).

The study day will be displayed in the data listings. If an event starts before the reference start date, the study day displayed on the listing will be negative.

Time unit

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

2.1.1.7 Baseline

Groups 1 and 2: For efficacy evaluations and PROs, randomization baseline will be used:

- The last non-missing assessment, including unscheduled assessments on or before the randomization date is taken as "baseline" value or "baseline" assessment.

Groups 1 and 2: For safety evaluations:

- In the placebo-controlled period, the last available assessment including unscheduled assessments on or before the *date of start of study treatment* (alpelisib or placebo) is taken as "baseline" assessment.
- In the Alpelisib period, the last available assessment including unscheduled assessments on or before the *date of start of investigational drug* (alpelisib) is taken as "baseline" assessment.

Groups 4 and 5: For safety and efficacy evaluations, the last available assessment on or before the *date of start of study treatment* (alpelisib) is defined as "baseline" assessment.

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

2.1.1.8 On-treatment assessment/event and observation periods

In general, for safety reporting, the observation period will be divided into three mutually exclusive segments:

1. Pre-treatment period: from day of participant's informed consent to the day before first dose of study treatment
2. On-treatment period: from day of first dose of study treatment to 30 days after last dose of study treatment
3. Post-treatment period: starting at day 31 after last dose of study treatment

In case at time of the analysis the date of last administration of study treatment is missing, on-treatment adverse event/safety assessment will include any adverse event/safety assessment recorded in the database and which occur after the start date of the study treatment.

The data summarized during the on-treatment period for the placebo-controlled period based on Table 2-1 will be as follows:

- For participants who discontinue treatment in either arm *prior to completing 16 weeks* of treatment, assessments/events with a start date from day of first dose of study treatment to 30 days after last dose of study treatment will be summarized
- For participants in the *placebo arm who switch to alpelisib after week 16*, only assessments/events with a start date prior to the first dose of alpelisib will be summarized
- For participants in the *alpelisib arm (who continue treatment after week 16)* assessments/events with a start date prior to the week 17 visit date/the first dose of open-label alpelisib will be summarized

The data summarized during the on-treatment period (for investigational drug alpelisib) for the Alpelisib period based on Table 2-1 will be as follows:

- For participants who discontinue treatment (at any time in *alpelisib arm*; from Week 17 Day 1 visit day/the first dose of open-label alpelisib for *placebo arm*), assessments/events with a start date from day of first dose of alpelisib up to 30 days after last dose of alpelisib will be summarized
- For participants in the *placebo arm who switch to alpelisib after week 16*, only assessments/events with a start date on or after the date of first dose of alpelisib will be summarized
- For participants randomized to the *alpelisib arm* assessments/events with a start date from first dose of alpelisib will be summarized

Similar conventions as for Group 2 will be used for the exploratory groups.

Data listings will include all assessments/events, flagging those which are not on-treatment assessment/event (i.e., pre-treatment and post-treatment period). In addition, flags for the analysis period (based on Table 2-1) or information on study period will be included in all listings.

2.1.1.9 Windows for multiple assessments

In order to summarize performance status, lab, vital signs, or other safety data collected over time (including unscheduled visits), the assessments will be allocated to a time window. The following general rule will be applied in creating the assessment windows: If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. If 2 assessments within a time window are equidistant from the target date, the selection should be made by selecting the one assessed by central (if any) and otherwise - for multiple central assessments equidistant to the planned visit - then the last/worst value will be used. If multiple assessments on the same date then the worst case will be used. Data from all assessments (scheduled and unscheduled), including multiple assessments, will be listed.

For instance, Table 2-2 shows time windows for Karnofsky/Lansky performance status assessments. The mid-point between the planned study day for two consecutive visits is taken

to define time windows. For assessments with a visit schedule different to the Karnofsky/Lansky performance status (e.g., ECGs, coagulation, ...), time windows based on mid-points will be defined in a similar manner. Refer to protocol Table 8-2 for the visit schedule for each assessment. For instance, HbA1c lab parameter is assessed at screening, weeks 12, 24, 36, 48, and thereafter every 24 weeks until week 260/End of treatment. The time window for the first scheduled assessment at Week 12 will range from Day 2 to 125.

Table 2-2 Time windows for Karnofsky/Lansky performance status assessments

Assessment	Target day of assessment	Time Interval
Baseline		≤ Day 1
Week 2 Day 1	8	Day 2 to day 18
Week 4	28	Day 19 to day 41
Week 8	56	Day 42 to day 69
Week 12	84	Day 70 to day 98
Week 17 Day 1	113	Day 99 to day 126
Week 20	140	Day 127 to day 153
Week 4*j (j in [6, 11])	d= j*28	Day d-14 to day d+13
Week 48	336	Day 322 to 419
Week 24*k (k ≥ 3]	d= k*6*28	Day d-84 to day d+83
End of Treatment		Assessment taken at the end of treatment visit

2.2 Analysis sets

The main efficacy and safety analyses will focus on the randomized participants in Groups 1 and 2. [redacted]

The **Full Analysis Set (FAS)** for Groups 1 and 2 comprises all participants to whom study treatment has been assigned by randomization. According to the intent to treat principle, participants will be analyzed according to the treatment and age group they have been assigned to during the randomization procedure.

The **Full Analysis Set – BYL719 (FAS – BYL719)** comprises all participants to whom alpelisib has been assigned by randomization. This analysis set will be used for the analysis of the primary endpoint.

The **FAS** for Groups 4 and 5 comprises all participants to whom study treatment has been assigned and who received at least one dose of study treatment. Participants will be analyzed according to the treatment they have been assigned to.

The **Safety Set** includes all participants who received at least one dose of study treatment. Participants will be analyzed according to the study treatment received. The actual treatment received corresponds to the randomized treatment if participants took at least one dose of that treatment, and to the first treatment received if the randomized treatment was never received.

For the analysis of Alpelisib period, the Safety set includes all participants who received at least one dose of investigational drug.

The **Pharmacokinetic analysis set (PAS)** includes all participants who receive at least one dose of alpelisib and provide at least one evaluable PK concentration. Participants will be analyzed according to the treatment received.

2.2.1 Participant classification

Participants may be excluded from the analysis sets defined above based on the protocol deviations entered in the database and/or specific participant classification rules as defined in Table 2-2.

Table 2-3 Participant classification based on protocol deviations and non-protocol deviations criteria

Analysis set	Protocol deviations leading to exclusion	Non-protocol deviation criteria leading to exclusion
FAS	No written informed consent	Groups 1 and 2: Not applicable Groups 4 and 5: No dose of study treatment
Safety Set	No written informed consent	No dose of study treatment
PAS	No written informed consent	No dose of study treatment or no evaluable PK concentration

2.2.2 Withdrawal of informed consent

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

Additional data for which there is a separate informed consent, e.g., [redacted] etc., collected in the clinical database without having obtained that consent will not be included in the analysis. These data will be excluded by the presence of the appropriate protocol deviation criterion.

2.2.3 Subgroup of interest

The following subgroups will be used for specific analyses, if there are enough participants in each category allowing such analysis:

- Age (6-11 years; 12-17 years)
- Gender
- Mutation type (as assessed locally; e.g., frequent, less frequent)
- PROS syndrome or phenotype (e.g., CLOVES, Klippel-Trenaunay syndrome (KTS))
- Lesion type (e.g., vascular, adipose)

Chinese subgroup analysis will be conducted to support China submission, and details will be elaborated in a separate SAP.

2.3 Participant disposition, demographics and other baseline characteristics

All demographic and other baseline characteristics including disease characteristics and participant characteristics will be summarized descriptively by age group and listed using the FAS. All listings and tables will be presented by age group and by treatment arm.

Descriptive analyses will be conducted where categorical data (e.g., gender, age categories, ethnicity, PROS syndrome, PIK3CA mutation) will be summarized by frequency counts and percentages; the number and percentage of participants with missing data at baseline will be provided. Continuous data (e.g., age, weight, height etc.) will be summarized by descriptive statistics (N, mean, median, standard deviation, minimum and maximum). For selected parameters 25th and 75th percentiles will also be presented.

Separate summaries will be presented for relevant medical histories and current medical conditions at baseline by age group. The summaries will be presented by primary system organ class (SOC) and preferred term (PT). Medical history, comorbidities and current medical conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology. The MedDRA version used for reporting will be specified in the CSR and as a footnote in the applicable tables/listings.

2.3.1 Participant disposition

Number (%) of participants screened will be summarized by country and center. In addition, the number (%) of participants randomized to study treatment/enrolled will be summarized by country, center and treatment arm. For participants who are screen failures, the reasons for not completing screening will be summarized.

The number (%) of participants in the FAS who were randomized/enrolled, who started treatment, are still on treatment, completed the study, discontinued from treatment, will be summarized together with the respective reasons for treatment discontinuation by analysis period and study period.

The following summaries will be provided (with % based on the total number of FAS participants) by age group and by treatment arm:

- Number (%) of participants who were randomized/enrolled *(based on data from IRT system)*
- Number (%) of participants who were randomized but not treated *(based on 'DAR' eCRF page not completed for any study treatment)*
- Primary reason for not being treated *(based on disposition eCRF page)*
- Number (%) of participants who were treated (*based on 'DAR' eCRF pages of each study treatment completed with non-zero dose administered*)

For each study treatment period (Core up to < Week 17 Day 1, Core Week 17 Day 1 to Week 24, Extension 1, Extension 2) and at any time:

- Number (%) of participants who are still on-treatment (*based on the disposition page not completed*);

- Number (%) of participants who discontinued the study treatment period *(based on the disposition page*)
- Primary reason for study treatment period discontinuation (*based on the disposition page*)
- Number (%) of participants who completed the study treatment period *(based on the disposition page)*
- Number (%) of participants who completed the study treatment period and did not enter the subsequent study treatment period

All disposition information will be listed.

Protocol deviations

The impact of COVID-19 on protocol adherence will be captured directly in the Study Visit, Dosing and Disposition domains by including the relationship as per the categories below. For all important protocol deviations, the relationship will also be captured.

The relationship to COVID-19 is defined using the following descriptions:

- COVID-19 Health Status: (i.e., participant's COVID-19 infection led to this PD)
- COVID-19 Site issue: (e.g., site closed, personnel not available)
- COVID-19 Lockdown: (e.g., site is active but patient not allowed to come)
- COVID-19 Subject/Patient concern: (e.g., site is active, subject/patient could come but refused to come / do assessment)
- COVID-19 Drug supply issue (e.g., drug was delivered to home)
- COVID-19 Other: (e.g., situation not already covered by the information above)

The number (%) of participants in the FAS with any protocol deviation will be tabulated by deviation category and summarized by age group, treatment arm. PDs that are not pandemic related and those that are related to the pandemic will be summarized separately, and in total. All pandemic related PDs will also be summarized by relationship to COVID-19.

All protocol deviations will be listed (with COVID-19 relationship as applicable).

Analysis sets

The number (%) of participants in each analysis set will be summarized by treatment arm and age group for the FAS. A listing will be provided displaying all participants excluded from analysis sets.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

Safety set will be used for the analyses described below. All listings and tables will be presented by age group and treatment arm.

Similar conventions as for Group 2 will be used for the exploratory groups.

2.4.1 Study treatment / compliance

Duration of exposure

The duration of exposure (in weeks) to alpelisib and to placebo will be summarized based on the analysis periods (i.e. placebo-controlled period and Alpelisib period). Summary statistics and categorical analyses (e.g., exposure <16 weeks, 16 - <24 weeks, 24 - <48 weeks, 48 - < 72 weeks etc.) will be presented. Details on start and end dates used for derivations are provided in Section 2.1.

Placebo-controlled period:

- Duration of exposure (days) = (min [date of last administration of blinded study treatment]) – (date of first administration of study treatment) + 1

Alpelisib period:

- Duration of exposure (days) = (min [last administration of investigational drug, data cut-off date]) – (date of first administration of investigational drug) + 1

In the Alpelisib period, exposure to alpelisib in Groups 1 and 2 will be presented overall and separately for participants in alpelisib arm and for participants in the placebo arm.

The above duration calculation includes the periods of temporary interruptions.

Duration of exposure may be graphically displayed.

Dose intensity

The actual dose intensity (computed as the ratio of actual cumulative dose received and actual duration of exposure) and the relative dose intensity (computed as the ratio of actual dose intensity and planned dose intensity) will be summarized for each study treatment by descriptive statistics.

The actual cumulative dose in mg is the sum of "dose administered" from the eCRF during the exposure of alpelisib/placebo.

Dose intensity will be summarized based on the analysis periods (i.e., placebo-controlled period and Alpelisib period).

Dose changes, interruptions or permanent discontinuations

The number (%) of participants with any dose changes (i.e., increases, reductions, interruptions, or permanent discontinuations) and the reasons (e.g., AE, dosing error, dispensing error, physician decision) will be taken from the 'Study Treatment eCRF' and summarized.

An **interruption** is defined as a zero dose on one or more days between two non-zero doses. The total duration of interruptions by participant will be summarized for the study population by time intervals, e.g. <1week, ≥1-<2 weeks, ≥2-<3 week etc. (these time intervals may be adjusted depending on the observed data). For participants who permanently discontinued study drug and the last dose record is 0mg, then the last zero dose will not be considered a dose-interruption.

A **dose increase** is defined as an increase as compared to the starting dose level of the period in question. Note that a dose rechallenge e.g. when a participant goes back to the previous prescribed dose after an interruption, is not considered as an increase (e.g., the sequence 50mg daily – 100mg daily – 0 mg – 50mg daily – 100mg is considered a dose rechallenge). If there is a dose increase due to a dosing error, then this is considered as **overdose**.

A **dose reduction** is defined as a decrease in dose from the starting dose (e.g. from 125mg daily to 50mg daily) even if the dose decrease has been directly preceded by an interruption (e.g., 125mg daily – 0mg – 50mg daily). On the other hand, if the dose decrease is followed by an interruption with the dose resuming at the same level prior to the interruption (e.g. in the sequence 125mg daily - 0 mg - 125mg daily), the dose decrease or change in dosing frequency will not be counted as dose reduction

All dose adjustments, interruptions and discontinuations will be summarized based on the analysis periods (i.e., placebo-controlled period and Alpelisib period). Dose reductions, interruptions and discontinuations may be graphically displayed using heatmaps.

2.4.2 Prior and concomitant therapies

Concomitant medications prior to and after the start of the study treatment will be coded using the World Health Organization (WHO) Drug Reference Listing (DRL) dictionary Anatomical Therapeutic Chemical (ATC) classification system and summarized by age group, treatment arm, lowest ATC class and preferred term within each analysis period using frequency counts and percentages.

Concomitant surgical and medical procedures not related to PROS will be summarized by age group, treatment arm, system organ class and preferred terms using frequency counts and percentages.

Prior and concomitant surgeries/invasive procedures for PROS will be summarized separately by age group, treatment arm and analysis period.

Separate summaries for pain medication and coagulopathy (see Section 5.4.3) will be presented by age group, and treatment arm.

2.5 Analysis of the primary estimand(s)

The primary objective is to demonstrate the efficacy of alpelisib based on the proportion of participants randomized to alpelisib with a confirmed objective response by BIRC in the pediatric (6-17 years old; Group 2) or in the adult (≥ 18 years old; Group 1) groups (multiple primary endpoints).

2.5.1 Definition of primary estimand(s)

The primary efficacy variable of the study is the proportion of participants randomized to alpelisib achieving confirmed objective response, where response is based on

- achieving ≥ 20% reduction from baseline in the sum of target lesion volumes (1 to 3 target lesions, assessed by MRI by BIRC),
- provided that none of the individual target lesions has ≥ 20% increase from baseline and
- in absence of progression of non-target lesions and without new lesions.

- Confirmation of response requires a subsequent imaging assessment performed at least 4 weeks after the onset of the response.

Refer to Section 2.5.3 for handling of intercurrent events and Section 2.5.4 for handling of missing values. Refer to Section 5.1.3 for the handling of response assessment dates in case of multiple imaging dates are assigned to one visit.

Analysis of the primary endpoint will be performed in each of Groups 1 and 2.

The percent change from baseline in the sum of target lesion volumes for each participant at Week XX, will be calculated using the following equation:

$$\frac{Sum(volume1_{XX\ weeks},\ volume2_{XX\ weeks}, volume3_{XX\ weeks}) - Sum(volume1_{baseline},\ volume2_{baseline}, volume3_{baseline})}{Sum(volume1_{baseline},\ volume2_{baseline}, volume3_{baseline})} \times 100\%$$

2.5.2 Statistical hypothesis, model, and method of analysis

The primary analysis will be based on data from the FAS - BYL719 for Groups 1 and 2 respectively. The study will be declared positive if at least one of the two primary null hypotheses can be rejected.

In order to conserve the overall type-1 error (one-sided level of significance of α=0.025) in testing multiple hypotheses, an alpha split with a graphical gate-keeping approach will be implemented based on the graphical multiple testing procedure as described in Section 2.6.1.

A confirmed response rate of 15% or less is considered "as insufficient" level of efficacy for the proposed participant population. Please refer to details in Global Protocol Amendment rationale 01 for the threshold for futility (updated from 35% to 15%). For the primary analysis, the following statistical hypotheses will be tested based on the exact binomial distribution:

- H_1 (Group 1): H_{01}: $p_{BYL} \leq 0.15$ vs. H_{A1}: $p_{BYL} > 0.15$
- H_2 (Group 2): H_{02}: $p_{BYL} \leq 0.15$ vs. H_{A2}: $p_{BYL} > 0.15$

where p_{BYL} is the confirmed objective response rate on alpelisib in the Group 1 and 2, respectively.

The primary null hypothesis will be rejected based on the probability of obtaining the observed response rate under a binomial distribution with underlying parameter p_0 =0.15 at the appropriate α-level governed by the graphical gatekeeping procedure described in Section 2.6.1 and Figure 2-1. With a total of 52 participants, if 15 or more confirmed responses are observed (observed confirmed objective response rate of 28.8%, lower bound of the 97.5% CI exceeding 15%), the null hypothesis will be rejected at a one-sided α/2= 0.0125.

Confirmed objective response rates in Groups 1 and 2 will be summarized using descriptive statistics (N, %) along with 2-sided exact 100(1-2α_i)% confidence interval (CI) (Clopper-Pearson exact method), where α_i corresponds to the appropriate α-level adjusted for multiple testing.

2.5.3 Handling of remaining intercurrent events of primary estimand

The primary analysis will account for different intercurrent events as explained in the following:

a. participants discontinuing treatment prior to confirmation of response: Classified as a non-responder
b. participants receiving surgery as a rescue therapy for any PROS lesions (target or other) prior to confirmation of response: Classified as a non-responder

2.5.4 Handling of missing values not related to intercurrent event

The primary efficacy variable is confirmed objective response. Confirmed response is defined as two responses reported from two subsequent imaging assessments performed at least 4 weeks apart.

Missing or non-evaluable assessment(s) between two available response assessments does not preclude the participant from being a confirmed responder.

2.5.5 Supplementary analyses

The primary endpoint will be summarized and reported for subgroups defined in Section 2.2.3, if there are enough participants in each category allowing such analysis.

Forest plots (response rate, 95% CI) will be produced to graphically depict the treatment effect estimates in different subgroups.

2.5.6 Sensitivity analyses for primary endpoint/estimand

As additional supportive analysis, the primary endpoint will be summarized based on the initial cut-off date, i.e. when all participants in Groups 1 and 2 have completed at least 24 weeks of treatment or discontinued earlier.

Sensitivity analysis may include potential analyses to assess the

- impact of COVID-19 pandemic
- the robustness of the volumetric lesion assessments for small lesions depending on the number of MRI slices a PROS lesion is visualized on

as applicable. Other sensitivity analyses may be done as needed.

2.6 Analysis of the key secondary endpoints/estimands

2.6.1 Key secondary estimand(s)

The analysis of secondary endpoints will be performed by age group and treatment arm. Periods defined based on pre- and post- switch of treatment will be considered for the analyses.

The key secondary objective is to demonstrate the efficacy of alpelisib based on the comparison of the proportion of participants achieving response at Week 16 with alpelisib versus placebo in Groups 1 or 2.

The key secondary analysis will be based on data from the FAS for Groups 1 and 2 respectively.

Definition of key secondary estimand(s)

A similar definition of response described for the primary efficacy variable in Section 2.5.1 will be used for the key secondary efficacy variable, using response at Week 16 instead of confirmed objective response. Participants who had a missing/non-evaluable radiological assessment at Week 16 will be considered as non-responders for the calculation of the key secondary efficacy variable. Handling of intercurrent events will be done as described for the primary estimand in Section 2.5.3: Instead of confirmed objective response for the primary, Week 16 will be considered for the key secondary estimands.

Statistical model, hypothesis, and method of analysis

The following statistical hypotheses will be tested

- H_3 (Group 1): H_{03}: $p_{BYL,W16} \leq p_{PBO,W16}$ *vs.* H_{A3}: $p_{BYL,W16} > p_{PBO,W16}$
- H_4 (Group 2): H_{04}: $p_{BYL,W16} \leq p_{PBO,W16}$ *vs.* H_{A4}: $p_{BYL,W16} > p_{PBO,W16}$

where $p_{BYL,W16}$ is the response rate on alpelisib at Week 16 and $p_{PBO,W16}$ is the response rate on placebo at Week 16. The analysis to test these hypotheses will consist of a Fisher's exact test at the appropriate α-level governed by the graphical gatekeeping procedure described below and in Figure 2-1.

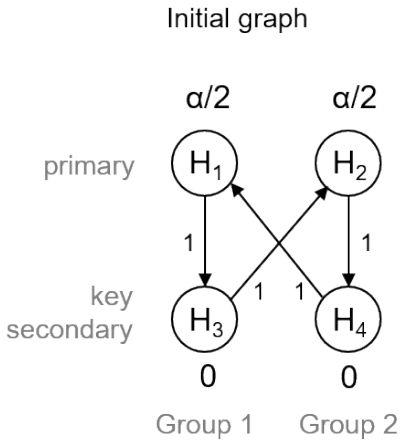
Response rates at Week 16 in Groups 1 and 2 will be summarized using descriptive statistics (N, %) along with 2-sided exact 95% CI (Clopper-Pearson exact method). The difference between treatment arms (alpelisib - placebo) in response rates at Week 16 will be presented together with 2-sided 95% CI (Santner-Snell exact method; unadjusted for multiple testing), separately for Groups 1 and 2.

Sequentially rejective graphical procedure for multiple testing

The 4 null hypotheses (primary hypotheses H_1 and H_2 described in Section 2.5.2; key secondary hypotheses H_3 and H_4 described above) are tested using the graphical approach to sequentially rejective multiple test procedures proposed by Bretz et al 2009. Significance levels α_i, i=1,…,4, are initially defined such that they sum up to α. The 2 primary hypotheses H_1 and H_2 are allocated with the levels $\alpha_i = 1/2\alpha$, i=1 or 2, where α = 0.025 (one-sided), i.e. the two primary hypothesis are considered equally important. For the 2 key secondary hypotheses H_3 and H_4, $\alpha_i = 0$, i=3 or 4, i.e. the key secondary hypothesis is not tested until its parent primary hypothesis is rejected.

The procedure then is as follows: Test the hypotheses H_i, i=1,…,4, each at its local significance level α_i. If a hypothesis H_i is rejected, remove H_i from the graph and propagate its level to other hypotheses according to a pre-specified rule represented by a directed, weighted graph. Update the reduced graph and repeat the testing step for the remaining, non-rejected hypotheses with the updated local significance levels. The procedure is repeated until no further hypothesis can be rejected. The procedure is fully determined by the initial graph given below and an updating algorithm, see Bretz et al 2009.

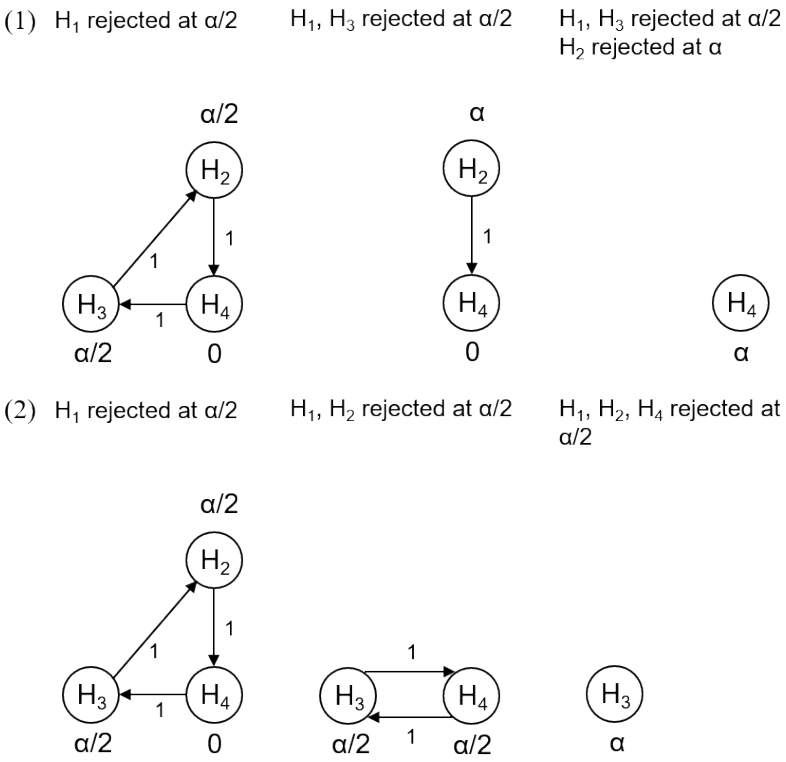
Figure 2-1 Graphical gate-keeping procedure to test primary and key secondary endpoints in order to control overall type-1 error



Edges with initial weights 0 (e.g. from H4 to H3) are not displayed in the graph. Initial significance levels and initial graph will be updated as per the updating algorithm in Bretz et al (2009).

This particular gatekeeping procedure can be interpreted as a Bonferroni-Holm test applied to the hierarchical pairs of primary and key secondary hypotheses (H_1, H_3) and (H_2, H_4) (Maurer et al 2011). To illustrate the sequentially rejective graphical procedure, Figure 2-2 displays two example rejection sequences. Note that in case more than one hypothesis could be rejected at a particular step, the algorithm guarantees that the sequence of rejection has no influence on the test decision (Bretz et al 2009). For instance, if one of the primary hypotheses can be rejected at α/2 (say Group 1), this level is first shifted entirely to its descendant key secondary hypothesis. If this can be rejected at level α/2, its level is shifted and added to the level of the remaining, nonparent primary hypothesis (Group 2) which in turn now can be tested at full level α. After the potential rejection of this second primary hypothesis, its descendant secondary hypothesis then can be tested at level α (example 1 in Figure 2-2).

Figure 2-2 Example rejection sequences for the sequentially rejective graphical procedure



Two example rejection sequences based on the initial graph in Figure 2-1.

Sensitivity and supplementary analyses for the key secondary estimands

The key secondary endpoint will be summarized and reported for the subgroups described in Section 2.2.1, if there is enough participants in each category allowing such analysis.

2.7 Analysis of secondary efficacy objective(s)

The FAS will be used for all analyses in this section unless otherwise stated. All listings and tables will be presented by age group and treatment arm.

In addition, for the primary analysis CSR data from participants who start alpelisib after completing 16 weeks of treatment on placebo will be flagged in listings.

2.7.1 Changes in PROS lesions

Changes in PROS lesion volume (by BIRC) over time will be reported in the following ways

- The proportion of participants with response at each scheduled assessment visit.
- The actual and percentage change from baseline in the sum of target lesion volumes,
- The actual and percentage change from baseline in the sum of MRI-measurable non-target lesion volumes
- The actual and percentage change from baseline in the sum of all MRI-measurable lesion (target and non-target) volumes

Response (by BIRC) at any assessment visit (Week XX) is defined as:

- achieving ≥ 20% reduction from baseline in the sum of target lesion volumes (1 to 3 target lesions, via BIRC) at Week XX,
- provided that none of the individual target has ≥ 20% increase from baseline at Week XX and
- in absence of progression of non-target lesions and without new lesions

Response (as per the criteria above) at a given time will be considered in the analysis only in the absence of progression. Handling of intercurrent events and of missing values will be done similar to Sections 2.5.3 and 2.5.4 for the primary estimand.

Waterfall graphs will be used to display the percentage change from baseline in the sum of target lesion volumes and as appropriate all MRI-measurable lesion (target and non-target) volumes for each participant at Week 16 and at Week 24. Best percentage change may also be displayed in a waterfall graph. The proportions of participants with various degrees of changes in the target lesion volume can be read directly from the graph. Only participants with both baseline and post baseline assessment at the timepoint of interest will be included in the waterfall graphs. Plots of PROS lesions volumes over time will be displayed.

Changes in other non-target lesions (by BIRC) will be reported in the following ways

- The proportion of participants with changes in non-target lesions i.e. number and percentage of participants evaluated with an improvement, stability or progression in the overall level of non-target disease burden.
- The proportion of participants with new lesions i.e. number and percentage of participants with a new lesion identified at each time point

2.7.2 Changes in symptoms and complications/comorbidities associated with PROS

Changes in symptoms and complications/comorbidities (i.e., PROS-related medical conditions, walking impairment, cardiac/pulmonary/renal function, pain, Karnofsky/Lansky performance status) will be summarized descriptively by age group. For each complication/comorbidity, only participants who have this reported as being present at baseline will be included in the analysis. When new symptoms/complications/comorbidities appear at any time of study therapy, they will be assessed for clinical significance and reported as adverse events when applicable.

The shift in CTCAE grade from baseline to selected key time points post-baseline (e.g., Week 16, Week 24, Week 48, etc.) for PROS-related conditions ongoing at baseline will summarized by preferred term and graphically displayed by age group and treatment. Improvement in PROS conditions is defined based on at least one grade reduction or resolution of the event. Worsening in PROS conditions is defined based on at least one grade increase of the event.

Changes from baseline at Week 24 will be presented by age group for the alpelisib arm using the FAS-BYL719 analysis set.

This analysis will be repeated on the FAS and changes from baseline at Week 16 will be presented for the placebo and alpelisib treatment arms by age group.

Note: The status at baseline will be summarized descriptively. The analyses of change from baseline in patients with impairment at baseline will be done as applicable if data from enough participants are available.

- Assessment of mobility will be performed at screening and then for those participants who have PROS-related impairment of mobility at efficacy assessment time points. Total distance covered by the 2 Min Walk Test (2MWT) and information on the participants need to use additional supporting tools to enable mobility (for instance, walking stick, moving chair), will be collected when applicable. The total distance walked during the 2MWT will be summarized by descriptive statistics at baseline and change from baseline will be presented over time. Any walking aids used for the test will be summarized and listed as appropriate.
- The left ventricular ejection fraction (LVEF) will be evaluated by echocardiography at Screening. In participants evaluated with PROS-related clinically relevant decrease of LVEF at screening, an echocardiography will be performed at Week 16 and Week 24 in the core phase of the study and at scheduled visits during the extension periods. Summary statistics will be provided for the changes in the LVEF measurement at scheduled timepoints from baseline.
- Spirometry testing will be performed at screening to assess a participant pulmonary function. For participants with pulmonary function affected by PROS, spirometry testing will be performed at Week 16 and 24 during the core phase and at scheduled visits during the extension period. Descriptive statistics to summarize the diffusion capacity of lung for CO, FEV1 and FVC (absolute and % predicted) at baseline will be provided along with changes from baseline at the scheduled efficacy timepoints.
- Descriptive summaries for change in creatinine clearance (estimated using MDRD for participants ≥ 18 years of age and creatinine-based Bedside Schwartz GFR equation for participants <18 years of age) from baseline will be presented at Week 16 and Week 24.
- The performance status measured with use of Karnofsky/Lansky scores will be summarized descriptively at baseline along with the change from baseline at scheduled efficacy timepoints.
- Analysis of the pain intensity is described in Section 2.11 as part of the patient reported outcomes.

2.7.3 Healthcare visits/hospitalization due to PROS, surgeries to manage PROS

Healthcare visits/hospitalizations and surgeries related to PROS during on-treatment period, will be summarized by age group, treatment arm and analysis period in the following ways

- Proportion of participants with healthcare visit/hospitalized due to PROS, and the number and duration of hospitalizations required
- Proportion of participants requiring surgery due to PROS, and the number of surgeries
- The proportion of participants where an anticipated PROS surgery is avoided due to improvement.

2.7.4 Overall clinical response

The proportion of participants with overall clinical response reported as an improvement, stable or worsening of clinical condition compared to baseline, as assessed by the Investigator (see protocol section 8.3) at scheduled timepoints will be summarized descriptively by age group, treatment arm and analysis period.

2.7.5 Duration of response

Duration of response (DOR) is defined as the time from first documented response in participants who are on treatment with alpelisib until progression of PROS lesion (by BIRC) or death. The analysis will include only participants with a confirmed response (see Section 2.5.1)

The start date of the first documented response will be the assessment date by which a participant will achieve ≥ 20% reduction from baseline in the sum of target lesion volumes (1 to 3 lesions, via BIRC) for the first time (i.e. the first of the two assessments that contribute to the confirmed response), provided that none of the individual target lesions has ≥ 20% increase from baseline that timepoint and in absence of progression of non-target lesions and without new lesions.

The end date is defined as the date of the first documented progression of PROS lesions by BIRC or death. For participants receiving rescue surgery on any PROS lesion after achieving response, the date of surgery will be considered as the end date of the response. Participants continuing without an event will be censored at the date of last adequate PROS lesion assessment prior to cut-off.

Progression of PROS lesions will be recorded when at least one of the following criteria are met:

- Increase of ≥ 20 % of the volume of any individual target and/or MRI-measurable non-target PROS lesions relative to nadir (defined as best achieved response)
- Progression of other non-target PROS lesions
- Appearance of a new PROS lesion(s) (when the new lesion is at least 1 cm long and clearly attributed to PROS)

DOR will be listed and summarized for all participants in the FAS-BYL719 with a confirmed response. The distribution of DOR will be estimated using the Kaplan-Meier method and the median time to response will be presented along with 95% confidence interval only if a sufficient number of responses is observed.

2.7.6 Time to treatment failure

This analysis only applies to participants who are on treatment with alpelisib. Time to treatment failure (TTF) is defined as the time from randomization/alpelisib treatment start date until the discontinuation of study treatment due to lack of efficacy (including unsatisfactory therapeutic effect, disease progression) or safety reasons (including adverse events, death). Participants who complete the study or discontinue study treatment for other reasons (e.g. discontinuation due to Participant/Guardian decision, technical problems) will be censored at the date of last study treatment received.

TTF will be summarized for all participants in the FAS who received alpelisib. The distribution of TTF will be estimated using the Kaplan-Meier method and the median time to treatment failure will be presented along with 95% confidence interval only if a sufficient number of treatment failures is observed.

2.8 Safety analyses

For all safety analyses, the safety set will be used. All listings and tables will be presented by age group and treatment arm and analysis periods (placebo-controlled period and Alpelisib period) as appropriate. Safety summarized during the Alpelisib period focuses on safety of alpelisib treatment, safety of placebo arm is summarized during placebo-controlled period.

Safety summaries (tables, figures) include only data from the on-treatment period (defined in Section 2.1.1.8) with the exception of baseline data which will also be summarized where appropriate (e.g. change from baseline summaries). In particular, summary tables for adverse events (AEs) will summarize based only on-treatment events, with a start date during the on-treatment period (treatment-emergent AEs).

2.8.1 Adverse events (AEs)

All information obtained on adverse events will be displayed by age group and treatment arm.

AE summaries will include only AEs that started or worsened during the on-treatment period (i.e., the **treatment-emergent** AEs) by each analysis period. A **treatment-emergent AE (TEAE)** is defined as any AE that develops after initiation of the study treatments or any event already present before initiation of the study treatments that increases in severity or developed into SAEs following exposure to the study treatments. Events that are present at baseline and subsequently reduce in CTCAE grade will not be considered as a TEAE.

All AEs reported in the AE eCRF page will be listed along with the information collected on those AEs, e.g. toxicity grade, relationship to study drug, outcome, action taken etc. AEs that started during the pre-treatment and post-treatment period will be flagged. A separate listing including any suspected or confirmed SARS-Cov-2 infections will be provided. The search criteria used to identify SARS-CoV-2 infections will be defined in the eCRS (electronic Case Retrieval Strategy) for the compound.

AEs will be summarized by number and percentage of participants having at least one AE, having at least one AE in each primary system organ class (SOC) and for each preferred term (PT). A participant with multiple occurrences of an AE will be counted only once in the respective AE category. A participant with multiple CTCAE grades for the same preferred term will be summarized under the maximum CTCAE grade recorded for the event. AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.03.

In the AE summaries, the primary system organ class will be presented alphabetically, and the preferred terms will be sorted within primary SOC in descending frequency.

The following adverse event summaries will be produced selecting all or a subset of AEs depending on seriousness, relationship to alpelisib, outcome or action taken.

- All AEs (by SOC and by PT) and separately those considered related to study treatment.

- SAEs and separately those considered related to study treatment
- SAEs with number of occurrences (an occurrence is defined as >1 day between start and prior end date of record of same preferred term).
- Non-SAEs
- SAEs with fatal outcome and separately those considered related to study treatment
- AEs leading to discontinuation
- AEs leading to dose reduction or dose interruption
- AEs requiring additional therapy

All summaries will show 'All grades' (including AEs with missing grade) and 'Grades ≥ 3'.

Incidences of hyperglycaemia events will be graphically displayed for participants with a dose escalation.

2.8.1.1 Adverse events of special interest / grouping of AEs

An adverse event of special interest is a grouping of adverse events that are of scientific and medical concern specific to compound BYL719. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HGLTs (high level group terms), HLT (high level terms) and PTs (preferred terms). Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. A NMQ is a customized group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms and/or an existing SMQ, narrow or broad. These searches will be defined in the eCRS and a listing of search terms will be provided in the CSR. The most up-to-date version of the CRS will be used at the time of analysis.

For each specified AESI, number (%) of participants with at least one event of the AESI occurring during on treatment period will be summarized together with the individual preferred terms in that grouping. In addition, number (%) of participants with at least one AESIs by maximum CTC grade, related AESIs, serious AESIs as well as action taken and outcome of the respective AESI will be summarized.

2.8.1.2 Time to and duration of first occurrence

Any analyses of the time to onset of an AE or AESI, or of the duration of the first occurrence of an event will be done only if there is sufficient number of events for a meaningful analysis.

Time to first occurrence

Time to onset of grade 2 or worse toxicities will be summarized for the Alpelisib period for the following AESIs (based on latest available eCRS and considering the PT that occurs first only):

- Hyperglycemia
- Rash
- GI toxicity (Nausea, Vomiting, Diarrhea)

In addition, time (days) to onset of first grade 2 or worse Hyperglycaemia (based on laboratory data) will be summarized.

Depending on the observed number of events for other AESI (or AEs), the time to first occurrence of any CTC grade ≥ 2 AESI (or AEs) may be summarized.

Time to first occurrence of an event is defined as time from start of study treatment to the date of first occurrence of this event (or first event within an AE grouping), i.e. time in days is calculated as (start date of first occurrence of event) – (start of study treatment) +1.

In the absence of an event during the on-treatment period, the censoring date applied will be **the earliest** of the following dates:

- end date of on-treatment period (or the date of the last lab assessment on-treatment for lab data)
- data cut-off date
- death date
- withdrawal of informed consent date.

Failure curves (ascending Kaplan-Meier curves) will be constructed by age group and treatment arm. Median together with 95% confidence interval as well as 25th percentile and 75th percentile will be presented for each age group and treatment arm.

In addition, the median time to occurrence for the subset of participants who experienced the event of interest will be calculated. Simple descriptive statistics, median, min and max as well as 25th percentile and 75th percentile, will be presented.

The same analysis might be repeated for events of grade 3 and grade 4.

Duration of first occurrence

The duration of first occurrence of an event will be analyzed for the Alpelisib period for the following Preferred Term or lab parameters if there is sufficient number of events for a meaningful analysis:

- Grade 2 or worse Hyperglycaemia (laboratory data)

Depending on number of events, the duration of first occurrence of other events may be summarized.

Duration of first event is defined as time from onset to the date of resolution of the first event: (date of resolution of first event) – (date of onset of event) +1.

Resolution of an event means that there is an event / a lab value returning to grade ≤ 1 for participants with an increase in grading while on treatment.

Duration of first event will be presented for the subset of participants of the Safety set who experienced the event.

A participant will be censored for time to resolution, if there is no resolution during the on-treatment period. The same censoring rules as described in time to first occurrence apply for duration of first occurrence.

Failure curves (descending Kaplan-Meier curves will be constructed by age group and treatment arm. Median together with 95% confidence interval, as well as 25th percentile and 75th percentile, will be presented for each age group and treatment arm.

2.8.2 Deaths

Summaries for on-treatment deaths will be produced showing deaths reasons by SOC and preferred term. All deaths will be listed, pre and post treatment deaths will be flagged.

2.8.3 Laboratory data

Laboratory data will be listed by age group, treatment arm, participant, and visit/time and if normal ranges are available abnormalities will be flagged. Shift tables using the low/normal/high/ (low and high) classification will be used to compare baseline to the worst on-treatment value.

Grading of laboratory values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 4.03. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account.

CTCAE Grade 0 will be assigned for all non-missing values not graded as 1 or higher. Grade 5 will not be used.

For laboratory tests where grades are not defined by CTCAE v4.03, results will be categorized as low/normal/high based on laboratory normal ranges.

The following summaries will be generated separately for hematology, and biochemistry tests:

- Listing of all laboratory data with values flagged to show the corresponding CTCAE v4.03 grades if applicable and the classifications relative to the laboratory normal ranges

For laboratory tests where grades are defined by CTCAE v4.03:

- Worst post-baseline CTCAE grade (regardless of the baseline status). Each participant will be counted only once for the worst grade observed post-baseline.
- Shift tables using CTCAE v4.03 grades to compare baseline to the worst on-treatment value

For laboratory tests where grades are not defined by CTCAE v4.03:

- Shift tables using the low/normal/high/ (low and high) classification to compare baseline to the worst on-treatment value.

Change from baseline in selected laboratory tests (e.g. FPG, HbA1c) may also be summarized by visit.

Liver function parameters such as total bilirubin (TBL), ALT, AST and alkaline phosphatase (ALP) will be summarized. The number (%) of participants with worst post-baseline values as per Novartis Liver Toxicity guidelines will be summarized (refer to Section 5.3.3).

Specific analyses to assess the impact on renal function (as part of efficacy) are described in Section 2.7.2.

2.8.4 Other safety data

2.8.4.1 ECG and cardiac imaging data

Data handling

For ECG replicates, the average of the ECG parameters at that assessment will be used in the analyses.

Data analysis

PR, QRS, QT, QTcF and HR intervals will be obtained from triplicate 12-lead ECGs for each participant during the study. ECG data will be read and interpreted locally.

Categorical analysis of QT/QTc interval data based on the number of participants meeting or exceeding predefined limits (refer to Section 5.4.1) in terms of absolute QT/QTc intervals and changes from baseline will be presented.

The number and percentage of participants with notable ECG values (refer to Section 5.4.1) will be presented by age group and overall and by treatment.

All ECG data will be listed by age group, treatment arm, participant and visit/time, notable values will be flagged. In the listing, the assessments collected during the post-treatment period will be flagged.

Summary statistics will be provided by age group, treatment arm and visit/time.

See Section 2.7.2 for analysis of cardiac imaging data (LVEF) as part of efficacy objectives.

2.8.4.2 Vital signs

Vital sign assessments are performed in order to characterize basic body function. The following parameters are collected: height (cm), weight (kg), body temperature (°C), pulse rate (beats per minute), systolic and diastolic blood pressure (mmHg).

Data handling

Vital signs collected on treatment will be summarized. Values measured outside of on treatment period will be flagged in the listings.

Data analysis

All vital signs data will be listed by age group, treatment arm, participant, and visit/time and if ranges are available, abnormalities will be flagged.

For analysis of vital signs the clinically notable vital sign criteria are provided in Section 5.4.2.

The number and percentage of participants with notable vital sign values (high/low) will be presented by age group and by treatment arm.

Descriptive statistics will be tabulated by visit for the age groups and by treatment arm using absolute change from baseline values for each vital sign measure and weight.

2.8.4.3 Growth, bone/dental development and sexual maturation

Data on growth, bone/dental development and sexual maturation (Tanner stage) will be tabulated and listed for children and adolescents using the Safety Set.

Growth

Height and Body Mass Index (BMI) will be summarized at 6-month intervals, using the standard deviation scores (SDS, also called z-score), velocity and velocity SDS. SDS will be calculated based on published referenced height and BMI information (http://www.who.int/childgrowth/en/).

Height/weight/BMI SDS and height/weight velocity SDS will be summarized using descriptive statistics (mean, standard deviation, range) for each time window, as well as by presenting number of participants with SDS values lower/higher than 5th/95th percentiles respectively. All height/BMI SDS, velocity and velocity SDS data will be listed.

BMI (kg/m^2) will be calculated as weight [kg] / ($height^2$ [m^2]).

The Standard Deviation Score (SDS) or Z-scores are used to describe how far a measurement is from the median (average). The distribution of heights of all boys (or all girls) of a certain age follows a normal (or almost normal) distribution. The SDS of an observed point in this distribution will be calculated as follows:

$$SDS = \frac{observed\ value - median\ value\ of\ the\ reference\ population}{standard\ deviation\ value\ of\ reference\ population}$$

These scores are calculated differently for measurements that are distributed normally and non-normally in the reference population (WHO, 2008).

The height velocity which evaluates the rate of growth per year will be calculated on the difference with the previous measure of height (in cm) and the time period (in days) as:

Height velocity (cm / year) =

$$\frac{(height\ V_x - height\ V_{x-1}) * F}{date\ V_x - date\ V_{x-1}}$$

(with V_x = actual visit ; V_{x-1} = previous visit ; $F = 365.25$).

Weight velocity is calculated similarly. Velocity SDS is calculated as (velocity – mean) / SD, where mean and SD are obtained as the height-, weight-, sex- and age-specific values in Tables 5 to 8 in Baumgartner et al (1986), where the age category immediately above the participant's exact age (at the assessment date V_x) should be used.

Sexual maturation

Sexual maturation will be monitored by Tanner staging. Per EPIK-P2 protocol, Tanner stages are collected only for participants in Groups 2 and 5 (6-17 years of age at the time of informed consent).

Start of puberty: Start of puberty corresponds to attainment of Tanner stage 2 (pubic hair or breast/genitalia attains Tanner stage 2).

Precocious puberty: The definition of precocious puberty is start of puberty occurring before the age of 8 in girls and before the age of 9 in boys (Merckmanuals.com Precocious puberty and statural growth; Carel et al 2004, Midyett et al 2003).

Delayed puberty:

- Delayed puberty in girls is generally defined as the failure to attain Tanner Stage 2 (for both breast development and pubic hair) by ages reported in the literature as between 12 and 13 years of age, or the absence of menstruation by ages reported in the literature as between 15 and 16 years of age (or absence of menstruation within 5 years of attainment of Tanner Stage 2) (Fenichel 2012).
- Delayed puberty in boys is generally defined as the failure to attain Tanner Stage 2 (for both genitals and pubic hair) by 14 years of age (Palmert and Dunkel 2012).

The age at which Tanner Stages 2-5 are achieved by gender will be summarized descriptively. The summary statistics will be provided for the number and percentage of participants in the different Tanner stages and the number of participants with delayed puberty ongoing at baseline. All Tanner stage data will be listed. Participants who experience delayed or precocious puberty will be flagged in the listing.

Data of Tanner stages over time will be displayed graphically using heatmaps. The heatmap of Tanner stages will also include age at menarche for females.

Depending on number of participants, analyses of precocious and/or delayed puberty may be explored.

Precocious puberty:

Precocious puberty may be analyzed among participants in the safety set potentially at risk of precocious puberty at baseline. This includes participants who have not already started puberty (Tanner stage 1 for both pubic hair and breast or genitalia) at baseline and have not yet reached the age criteria for precocious puberty. In addition, participants within [6 to < 7] years for girls and within [6 to < 8] years for boys at baseline are considered potentially at risk during the study.

The number and percentage of participants with precocious puberty will be presented globally and for girls and boys separately by study. The number of participants for the denominator will be based on the participants at risk.

Delayed puberty:

Delayed puberty may be analyzed among participants in the safety set potentially at risk of delayed puberty at baseline. This excludes participants who were identified as having delayed puberty at baseline, male participants who already started puberty, and female participants who already started puberty and experienced menarche before baseline.

The number and percentages of participants with delayed puberty will be presented globally and for boys and girls separately by study. In case the follow-up period is not long enough to conclude on the puberty status, the status regarding delayed puberty could be undetermined for some participants. The number of participants for the denominator will be based on the participants at risk.

In addition, and if enough participants are available, the mean age reaching Tanner stage 2 for participants in Tanner stage 1 at baseline may be calculated for each gender.

Bone/dental development

Data on bone/dental development in children and adolescents (<18 years) will be summarized descriptively and listed.

Dental development will be assessed via routine method of visual evaluation. Panoramic radiographs may be performed as per local standards when clinically indicated. The proportion of participants with normal or abnormal results over time will be summarized or graphically displayed.

Bone age is assessed according to Greulich-Pyle standards (Greulich and Pyle, 1959, Sasaki et al 2002, and Hojreh et al 2018).

In order to take into account that bone age varies according to gender and age, bone age will be expressed in standard deviation scores (SDS). Bone age SDS will be calculated as (bone age – chronological age) / SD) where the chronological age is the age in months at the time of the imaging evaluation and SD is the sex- and age-specific standard deviation reported by Greulich and Pyle (1959).

Delayed or advanced bone age is defined as SDS < -2 or > 2 (values used by pediatric experts to identify potential abnormalities). This will allow an evaluation of the number and percentage of patients with advanced or delayed skeletal maturation in each treatment arm.

Bone age SDS will be summarized by treatment arm using descriptive statistics (mean, standard deviation, range) for each visit.

Individual trajectories for bone age will be displayed graphically on skeletal maturity charts as described in Boechat and Lee (2007).

2.8.4.4 Resource utilization

Data relating to resource utilization will be used for the purpose of economic evaluation and will be carried out and reported as a separate activity.

PROS-related healthcare visits/hospitalizations and surgeries will be analyzed as part of efficacy objectives.

2.9 Pharmacokinetic endpoints

PAS will be used in all pharmacokinetic data analysis and PK summary statistics.

Alpelisib plasma concentration data will be listed by age group/dose, participant, and visit/sampling time point. Descriptive summary statistics will be provided by visit/sampling time point, including the frequency (n, %) of concentrations below the LLOQ and reported as zero.

Summary statistics will include mean (arithmetic and geometric), SD, CV (arithmetic and geometric), median, minimum, and maximum. Concentrations below LLOQ will be treated as zero in summary statistics and for PK parameter calculations.

The pharmacokinetic parameters (Table 2-2) will be calculated from individual plasma concentration versus time profiles using non-compartmental analysis (Phoenix WinNonlin). Cmax and Ctrough are primary PK parameters; AUC0-24h, AUClast, AUCinf, Tmax, T1/2, CL/F and Vz/F may be evaluated as secondary PK parameters if data permit. Additional diagnostic pharmacokinetic parameters may be listed, including AUC%Extrap and Rsq_adj. Pharmacokinetic parameters will be listed by treatment and participant. Descriptive summary statistics will include mean (arithmetic and geometric), SD, and CV (arithmetic and geometric), median, minimum, and maximum. An exception to this is *Tmax* where median, minimum, and maximum will be presented.

Table 2-2 Pharmacokinetic parameters

AUC0-24h	The AUC from time zero to 24 hours after the dose (ng.hr/mL)
AUClast	The AUC from time zero to the last measurable concentration sampling time (Tlast) (ng.hr/mL)
AUCinf	The AUC from time zero to infinity (ng.hr/mL)
Tmax	The time to reach maximum (peak) plasma, blood, serum, or other body fluid drug concentration after single dose administration (hr)
Cmax	The observed maximum plasma (or serum or blood) concentration following drug administration (ng/mL)
Ctrough	The trough observed concentration is the concentration that is just prior to the beginning, or at the end, of a dosing interval (ng/mL). Synonymous with the concentration of “0H” PK sample.
T1/2	The elimination half-life associated with the terminal slope (λz) of a semi logarithmic concentration-time curve (hr).
CL/F	The total body clearance of drug from the plasma (L/hr)
Vz/F	The apparent volume of distribution during terminal phase (associated with λz) (L)
AUC%Extrap	AUC extrapolated from Tlast to infinity as a percentage of AUCinf
Rsq_adj	Square of the correlation coefficient (adjusted for the number of data points included) associated with λz

[redacted]

2.11 Patient-reported outcomes

The following Patient Reported Outcomes (PRO) tools will be evaluated using the FAS-BYL719 and FAS populations:

- Patient Global impression of symptom severity item
- Worst Pain intensity, item from the Brief Pain Inventory (BPI), as collected with the PRO Diary (for participants ≥12 years of age and older)
- Pain interference with general activity in the past 24 hours, as collected with the PRO Diary (for participants ≥12 years of age and older)
- Wong Baker FACES Pain Rating Scale (for participants 3-11 years)

- Location of pain, as collected with the PRO Diary
- Type of pain, as collected with the PRO Diary
- Dyspnea items, as collected with the PRO Diary
- PROMIS Profile domains

Age-appropriate versions of the PRO tools will be used for participants ≥ 18 years of age (in Group 1) and participants aged 12 to 17 years (in Group 2). Observer versions of PRO tools completed by parent/caregiver will be used for participants <12 years of age, except for Worst Pain intensity where participants will use Wong Baker Faces Pain rating scale, as able to report. Refer to protocol tables 8-7 and 8-8 for details. For the analysis, PRO scores based on adult self-reports (≥ 18 years old), child self-reports (≥12 to 17 years old), child reports with parent/caregiver assistance (3 to 11 years old), and parent/caregiver reports (<12 years old) will be presented separately.

The patient-reported global impression of symptom severity and BPI worst pain intensity item are identified as the primary PRO measures of interest.

Change from baseline in scores at Week 16 will be assessed on treatment with alpelisib as compared to placebo in pediatric (12 to 17 year old) and adult (≥ 18 year old) populations.

No multiplicity adjustment will be applied.

See Section 5.5.2 for the derivation of PRO total scores.

Patient Global impression of Symptom Severity (PGI-S)

A Patient Global Impression of Symptom Severity item will be used to understand the overall severity of symptoms experienced and clinical meaningfulness of treatment effects experienced during the study. This item includes 5 response options: no symptoms, mild, moderate, severe, and very severe. Minor modifications are made to the item wording to account for differences between adult/child self-report and parent/caregiver report (e.g., changing “your overall symptoms” to “your child’s overall symptoms”).

The number (%) of participants in each response category at baseline and at each scheduled timepoint will be summarized and presented graphically using barplots by age group and treatment arm.

At Week 16, the proportion of participants with a one category improvement from baseline in Group 1 (≥ 18 years of age) and in Group 2 (12 to 17 years of age) will be summarized using descriptive statistics along with 2-sided 95% CI (Clopper-Pearson exact method).

The difference in proportions between treatment arms (alpelisib - placebo) at Week 16 will be presented together with 2-sided 95% confidence interval by age group.

Shift tables to compare baseline to the last available assessment on treatment for the PGI-S item will be presented.

Percentage changes in target lesion volume will be summarized for participants with a 1 category improvement in the PGI-S scale at Week 16 and at Week 24. Data will be presented using cumulative distribution function (CDF) plots.

Evaluation of Pain intensity, location and type (collected with PRO diary)

The Brief Pain Inventory (BPI) is a multi-item assessment of pain; four items from this questionnaire will be used. The BPI item that assesses worst pain intensity in the past 24 hours will be used to assess pain intensity. Participants respond to the item on an 11-point response scale ranging from 0 (no pain) to 10 (pain as bad as you can imagine). The item that assesses interference with general activity in the past 24 hours will also be used. Participants respond to the item on an 11-point response scale ranging from 0 (does not interfere) to 10 (completely interferes). Additionally, the participant will indicate on the location of pain on a modified BPI diagram, a separate modified BPI item will allow the participant to indicate the type of pain experienced. Optional questions on worst pain intensity, location and type of pain will also be asked for an additional area of pain experienced by the participant, as applicable.

The Pain diary will be used to collect this information daily on the primary location of pain and a secondary location of pain along with the associated worst pain intensity and type of pain and pain interference as adapted from the Brief Pain Inventory (BPI). Worst pain intensity and pain interference will be averaged weekly over a 7-day period if a participant has completed the questionnaire for at least 4 days in the 7-day period. The weekly mean will be calculated based on the available assessments. See Section 5.5.3 for details.

Adult participants in Group 1 and pediatric participants 12 years of age and older in Group 2 that are able to understand this item will report pain with these questions.

In Group 2, for participants that are 11 year of age and younger, the Wong-Baker FACES Pain Rating Scale will be used. This scale is a single-item that includes simple line drawings of 11 faces, each depicting an increasing amount of pain. Each face is associated with both a numeric rating (ranging from 0-10) and a descriptor (ranging from no hurt – hurts worst). Children will be asked to choose the face that best describes their level of pain (Wong and Baker 1988).

For participants able to use the BPI questionnaire, the number (%) of participants will be summarized by the location of pain (primary and secondary) as well as the type of pain, weekly worst pain intensity, and weekly pain interference. Responses will be summarized at baseline and over time. Daily responses of worst pain intensity and pain interference will be presented graphically.

Clinically important change at Week 16 will be defined as a 2 point reduction for participants who have a pain intensity score ≥ 4 at baseline. For participants with a baseline score < 4, a 1 point reduction from baseline will also be considered as a clinically important change.

At Week 16, the proportion of participants with a clinically important change from baseline will be summarized using descriptive statistics along with 2-sided 95% CI (Clopper-Pearson exact method).

The difference in proportions between treatment arms (alpelisib - placebo) at Week 16 will be presented together with 2-sided 95% confidence interval by age group.

At Week 24, the proportion of participants with a clinically important change from baseline in pain scores will be summarized using descriptive statistics along with 2-sided 95% CI (Clopper-Pearson exact method).

Changes in target lesion volume will be summarized for participants with a clinically important change in the average pain score scale at Week 16 and at Week 24 and presented graphically using waterfall plots.

For participants using the Wong-Baker FACES Pain Rating Scale, the average weekly intensity of pain will be summarized.

Pain medications will be analyzed separately. Changes in pain medication (e.g., changes in use of opioids, non-opioid analgesics, no analgesics, see Section 5.4.3) will also be summarized descriptively/or graphically displayed for participants with and without a clinically important change in the average pain score scale at Week 16 and at Week 24.

Dyspnea items (collected with PRO diary)

Two items on Dyspnea Severity, assessed during cognitive debriefing interviews, will be included to assess the degree to which respondents felt short of breath while completing various physical activities. The two items selected from the item bank for inclusion are shortness of breath when walking up 10 stairs and shortness of breath when talking while walking, with a recall period of the past 24 hours. Each of the items includes 5 response options: no shortness of breath (0), mildly short of breath (1), moderately short of breath (2), severely short of breath (3), and I did not do this.

Dyspnea Severity score will be averaged weekly over a 7-day period if a participant has completed the questionnaire for at least 4 days in the 7-day period. The weekly mean will be calculated based on the available assessments and summarized by age group and treatment arm.

PROMIS Profile domains

In Group 1, the PROMIS-29 plus 2 Profile v2.1 (designed for adults ≥18 years of age) will be used to measure 29 items across 7 domains (4 items per domain): depression, anxiety, physical function, pain interference, fatigue, sleep disturbance, and ability to participant in social roles and activities. The 27th and 28th items are the “plus 2” and will be used to assess cognitive function abilities and the 29th item assesses pain intensity. All items include 5 response options, except for the pain intensity item, which has 11 response options (Cella et al 2019).

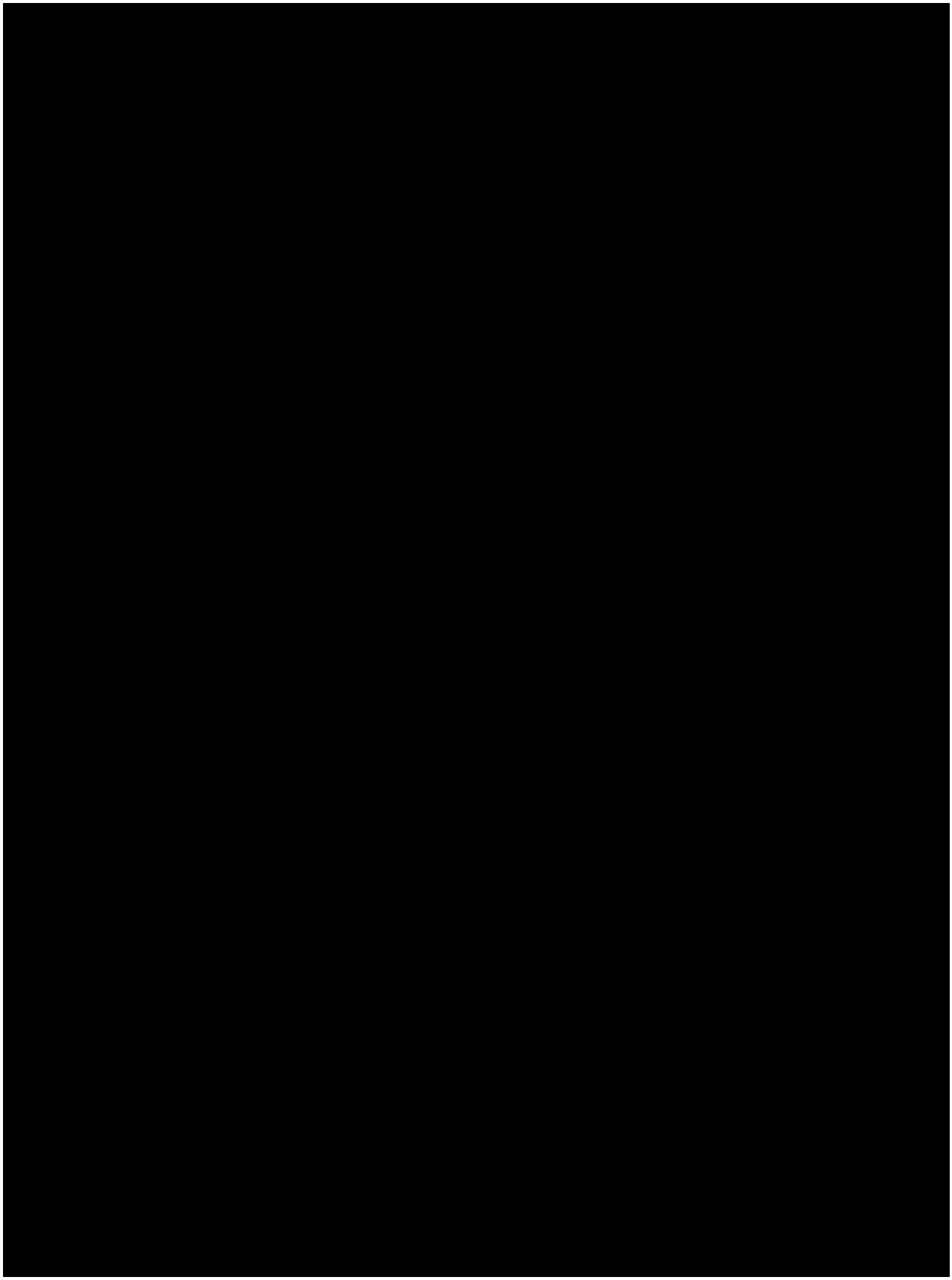
In Group 2, the PROMIS Pediatric-25 Profile v2.0 for self-reporting by children (12 years and older); the PROMIS Parent-Proxy-25 Profile v2.0 is designed for completion by a parent or caregiver observer on behalf of children (under 12 years of age) unable to record for themselves.. Four items are included for each of the following domains: depressive symptoms, anxiety, physical function-mobility, pain interference, fatigue, and peer relationships. The 25th item will assess pain intensity. All items include 5 response options, except for the pain intensity item, which has 11 response options.

The content of the child self-report and parent/caregiver versions are identical, except for minor modifications to account for differences in responders (e.g., “I felt worried” vs “My child felt worried”).

Unlike the adult PROMIS Profile, the pediatric and parent-proxy PROMIS profile does not include a sleep disturbance domain. A short form including 4 items that assess difficulty falling asleep, sleeping through the night, problems with sleep, and trouble sleeping will be used to assess the sleep disturbance domain. Each item is answered in reference to the past 7 days and includes 5 response options.

Each question usually has five response options ranging in value from one to five. All questions must be answered in order to produce a valid score using the scoring tables. Scoring of PROMIS data and methods for handling of missing items or missing assessments will be handled according to the scoring manual and user guide for each respective participant questionnaire (PROMIS scoring manual, also see Section 5.5.2).

Descriptive statistics will be used to summarize the standardized scores, as well as change from baseline, at each scheduled assessment timepoint. Participants with an evaluable baseline score and at least one evaluable post baseline score will be included in the change from baseline analyses. These scores will be also displayed graphically by scheduled assessment timepoint.



2.14 Interim analysis

There is no planned interim efficacy analysis before the conduct of the primary analysis.

3 Sample size calculation

3.1 Primary endpoint(s)

The sample size calculation is based on the proportion of participants with confirmed objective response (see Section 2.5.1 for the definition).

Based on the case series in Venot et al 2018 (derived from Supplementary Table 2), 11/17 (64.7%, 95% CI: 38.3 to 85.8%) of participants treated with alpelisib (50 mg for children, 250 mg for adults) had ≥ 20% reduction in volume of the target lesion as assessed by MRI by the Investigator at month 3. At month 6, 15/17 (88.2%, 95% CI: 63.6 to 98.5%) of participants had ≥ 20% reduction in volume. Considering only children treated with alpelisib 50 mg, 12/14 (85.7%, 95% CI: 57.2 to 98.2%) of participants had ≥ 20% reduction in volume at month 6.

In the retrospective EPIK-P1 study (CBYL719F12002), 12/32 (37.5%, 95% CI: 21.1 to 56.3%) of participants with PROS and treated with alpelisib were responders at Week 24 based on independent central radiological review and based on a complete case analysis (Canaud et al 2021). Up to 3 clinically relevant lesions were selected as target for each individual patient and assessed over time. Sensitivity analyses include one based on the efficacy analysis set (n=37) in which participants with missing response at 24 weeks or 6 months (+/- 4 weeks) were

considered as "non-responders" (i.e., 12/37=32.4% of participants, 95% CI: 18.0 to 49.8%). Across the 31 participants who had an imaging assessment at the index date and at Week 24, the reduction in target lesion volume was observed early and this improvement was sustained. All participants with a first response at or before Week 24 and subsequent MRI assessments continued to demonstrate a response. There were only 3 out of 37 patients (8.1%, 95% CI: 1.7 to 21.9%) with a radiological assessment at the index date who reached the criteria of response after Week 24 [CBYL719F12002 (EPIK-P1) CSR].

Based on the EPIK-P1 efficacy results, it is anticipated that most of the responses will be associated with an onset within the first 24 weeks of treatment and that responses will be subsequently sustained. Hence the expected confirmed response rate on alpelisib is anticipated to be similar to the expected response rate at Week 24, and sample size and power calculations in this section were done based on assumptions for response rate at Week 24.

In Adams et al 2016, 35% (95% CI: 20 to 51%) of participants with complicated vascular anomalies had ≥ 20% reduction in size (max diameter) of target lesion at 24 weeks of treatment with Sirolimus. In the Parker et al 2019 study of 39 participants with PROS and progressive overgrowth, a significant reduction of –7.2% (SD 16.0) was observed in the volume change of affected tissues at the end of 26 weeks Sirolimus treatment by DXA technique. Among the 22 participants with evaluable volume for affected tissue at the end of 26 weeks sirolimus treatment, none (0% of participants, 95% CI: 0 to 15.4%) achieved ≥ 20% reduction (derived from supplementary material Table S2 in Parker et al 2019).

It is expected that a spontaneous ≥ 20% reduction in lesion volume without treatment is rare.

Approximately 78 participants are planned to be randomized for each age group (Group 1 and Group 2) in two treatment arms in 2:1 ratio.

Based on the exact binomial distribution, approximately 52 participants are required in alpelisib arm of each age group to test a null hypothesis of confirmed response rate ≤ 15% vs. a target confirmed response rate of 35% or more, with a one-sided α of 1.25% and 86.0% power. For an observed confirmed response rate of 28.8% (15 responders in 52 participants), the exact 97.5% CI is 15.8 to 45.0% (Table 3-1). The probability for a successful study (i.e. rejecting at least one of the two primary null hypotheses) is 98.1%, assuming true confirmed response rates of 35% in both Group 1 and 2 (see Table 3-5). Further operating characteristics are provided in Section 3.2.

The key secondary objective is to compare the proportion of participants with response at Week 16 for alpelisib vs placebo in Group 1 or Group 2. The sample size (i.e. 78 participants in each group) was determined in order to obtain adequate operating characteristics as detailed in Section 3.2.

Therefore, a total of approximately 156 participants (of age ≥ 6 years) will be randomized in the study.

Exploratory Groups 3, 4 and 5

Two additional groups of approximately 18 participants (based on feasibility considerations) of age 2 to 5 years will be enrolled for exploratory purposes in Group 3 (12 participants, alpelisib granules formulation) and Group 4 (6 participants, alpelisib FCT formulation). The starting

dose (of the granules formulation) for Group 3 will be selected based on primary analysis from other age groups. Group 4 will be open to enrolment immediately after the implementation of Global Protocol Amendment 01, whereas Group 3 will be open to enrolment only after implementation of a future substantial Global Protocol Amendment.

Another group, Group 5, of approximately 15 participants of age 6 to 17 years will be enrolled for exploratory purposes. No formal sample size calculation has been performed. The precision of the estimate for the difference in confirmed response rates of alpelisib in Group 5 as compared to Group 2 is evaluated, assessing the width of 95% CIs (Santner-Snell exact method).

Table 3-1 Sample size based on exact binomial test for a testing response rate of alpelisib vs 15% (1.25% one-sided significance level)

Power	Assumed response alpelisib	N= Sample size per age group	R= Required number of responders to reject H0	Estimated response rate (97.5% CI) with sample size N and responders R
86.0%	35%	52	15	28.8% (15.8 to 45.0%)

Calculations were made using the software package PASS (11.0).

3.2 Secondary endpoint(s)

The key secondary objective is to compare the proportion of participants achieving response (see Section 2.6.1 for definition) at Week 16 with alpelisib versus placebo in Groups 1 or 2. Response rates at Week 16 will be compared between the two randomized treatment arms provided that the primary endpoint is statistically significant in the respective age group according to the graphical gatekeeping approach described in Section 2.6.1.

It is assumed that the response rate at Week 16 is 30% with alpelisib and 5% with placebo. Approximately 78 participants are to be randomized in the two treatment arms in 2:1 ratio to detect a difference in response rate of 30% vs 5% with 61.7% power using Fisher's exact test and one-sided α = 1.25%. (Table 3-4). Given the hierarchical testing strategy, this power is conditional on the primary endpoint being met in the respective age group.

Table 3-5 shows simulated probability for a successful study (at least one primary null hypothesis rejected) and individual powers for each of the 2 primary and 2 key secondary hypotheses based on the sequentially rejective graphical procedure with one-sided level of significance of α = 2.5% described in Section 2.6.1.

Table 3-4 Power based on Fisher's exact test for testing response rate at Week 16 of alpelisib vs placebo (1.25% one-sided significance level)

Power	N= Sample size per age group		Assumed response rate per age group	
	Alpelisib arm	Placebo arm	Alpelisib arm	Placebo arm
61.7%	52	26	30%	5%

Calculations were made using the software package PASS (11.0).

Table 3-5 Probability for a successful study (reject at least one of the 2 primary null hypotheses) and individual power for different scenarios

Assumed response rates										
Confirmed response on alpelisib		Response at Week 16 on alpelisib		Response at Week 16 on placebo		Prob. to reject H1 and/or H2	Individual power* to reject			
Age Gr. 1	Age Gr. 2	Age Gr. 1	Age Gr. 2	Age Gr. 1	Age Gr. 2		H1	H2	H3	H4
30%	35%	30%	30%	3%	3%	94.5%	68.8%	88.4%	55.0%	69.0%
35%	35%	30%	30%	3%	3%	98.1%	90.2%	90.2%	73.1%	73.3%
40%	35%	30%	30%	3%	3%	99.4%	97.5%	89.5%	78.1%	71.5%
30%	35%	30%	30%	5%	5%	94.1%	68.1%	87.2%	44.6%	55.9%
35%	**35%**	**30%**	**30%**	**5%**	**5%**	**98.1%**	**89.2%**	**89.3%**	**59.1%**	**58.9%**
40%	35%	30%	30%	5%	5%	99.4%	97.5%	88.2%	65.8%	59.9%

Calculations were made using the software package R (3.6.1). Power calculated based on 5000 simulated observations (from a binomial distribution with sample size per age group of N=52 on alpelisib, N=26 on placebo, and age-group specific response rate as parameters) for each scenario in the table.

The scenario in bold corresponds to the assumptions made for sample size calculation.

* Individual powers to reject H3 and H4 are marginal powers, accounting for the whole testing strategy.

4 Change to protocol specified analyses

There is no change from protocol specified analysis.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

The following rule should be used for the imputation of the dose end date for a given study treatment component:

Scenario 1: If the dose end date is completely missing and there is no EOT page and no death date, the participant is considered as on-going:

The participant should be treated as on-going and the cut-off date should be used as the dose end date.

Scenario 2: If the dose end date is completely or partially missing and the EOT page is available:

Case 1: The dose end date is completely missing, and the EOT completion date is complete, then this latter date should be used.

Case 2: Only Year(yyyy) of the dose end date is available and yyyy < the year of EOT date:

Use Dec31yyyy

Case 3: Only Year(yyyy) of the dose end date is available and yyyy = the year of EOT date:

Use EOT date

Case 4: Both Year(yyyy) and Month (mm) are available for dose end date, and yyyy = the year of EOT date and mm < the month of EOT date:

Use last day of the Month (mm)

Case 5: Both Year(yyyy) and Month (mm) are available for dose end date, and yyyy = the year of EOT date and mm = the month of EOT date:

Use EOT date

All other cases should be considered as a data issue and the statistician should contact the data manager of the study.

After imputation, compare the imputed date with start date of treatment, if the imputed date is < start date of treatment:

Use the treatment start date

Participants with missing start dates are to be considered missing for all study treatment component related calculations and no imputation will be made. If start date is missing then end-date should not be imputed.

5.1.2 AE, concomitant medication, medical history and safety assessment date imputation

Table 5-1 Imputation of start dates (AE, CM, MH, FAMH) and assessments (LB, EG, VS)

Missing Element	Rule
day, month, and year	• No imputation will be done for completely missing dates
day, month	• If available year = year of study treatment start date then o If stop date contains a full date and stop date is earlier than study treatment start date then set start date = 01JanYYYY o Else set start date = study treatment start date. • If available year > year of study treatment start date then 01JanYYYY • If available year < year of study treatment start date then 01JulYYYY
day	• If available month and year = month and year of study treatment start date then o If stop date contains a full date and stop date is earlier than study treatment start date then set start date= 01MONYYYY. o Else set start date = study treatment start date. • If available month and year > month and year of study treatment start date then 01MONYYYY • If available month and year < month year of study treatment start date then 15MONYYYY

Table 5-2 Imputation of end dates (AE, CM, MH)

Missing Element	Rule (*= earliest between (last treatment date plus 30 days, death date, cut-off date, withdrawal of consent date))
day, month, and year	• Completely missing end dates (incl. ongoing events) will be imputed by the end date of the on-treatment period*
day, month	• If partial end date contains year only, set end date = earliest of 31DecYYYY or end date of the on-treatment period *
day	• If partial end date contains month and year, set end date = earliest of last day of the month or end date of the on-treatment period*

Any AEs and ConMeds with partial/missing dates will be displayed as such in the data listings.

Any AEs and ConMeds which are continuing as per data cut-off will be shown as ‘ongoing’ rather than the end date provided.

5.1.3 Response/PROS lesion assessment dates

All investigation dates (e.g. MRI scan without contrast) should be completed with day, month and year. If one or more assessment dates are incomplete but other investigation dates are available, this/these incomplete date(s) are not considered for calculation of the assessment date and assessment date is calculated as the latest of all investigation dates (e.g. MRI scan without contrast) if the response at that assessment is "Response" or "Non-response-Non-PD" or "Unknown". Otherwise – if response is "Non-response-Progressive Disease" – the assessment date is calculated as the earliest date of all investigation dates at that visit.

5.2 AEs coding/grading

Adverse events are coded using the Medical dictionary for regulatory activities (MedDRA) terminology.

AEs will be assessed 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.

If CTCAE grading does not exist for an adverse event, grades 1-4 corresponding to severity of mild, moderate, severe, and life-threatening will be used.

5.3 Laboratory parameters derivations

5.3.1 Grading

Grade categorization of lab values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 4.03. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. The criteria to assign CTCAE grades are given in Novartis internal criteria for CTCAE grading of laboratory parameters.

For laboratory tests where grades are not defined by CTCAE v4.03, results will be graded by the low/normal/high (or other project-specific ranges, if more suitable) classifications based on laboratory normal ranges.

A severity grade of 0 will be assigned for all non-missing lab values not graded as 1 or higher. Grade 5 will not be used. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia.

5.3.2 Imputation Rules

CTC grading for blood differentials is based on absolute values. However, this data may not be reported as absolute counts but rather as percentage of WBC.

If laboratory values are provided as '<X' (i.e. below limit of detection) or '>X', prior to conversion of laboratory values to SI unit, these numeric values are set to X.

The following rules will be applied to derive the WBC differential counts when only percentages are available for a xxx differential

xxx count = (WBC count) * (xxx %value / 100)

Further derivation of laboratory parameters might be required for CTCAE grading. For instance, corrected calcium can be derived using the reported total calcium value and albumin at the same assessment using the following formula:

Corrected Calcium (mg/dL) = Calcium (mg/dL) – 0.8 [Albumin (g/dL)-4]

In order to apply the above formula, albumin values in g/L will be converted to g/dL by multiplying by 0.1), calcium values in mmol/L will be converted to mg/dL by dividing by 0.2495. For calculation of laboratory CTC grades 0 and 1, the normal range for derived corrected calcium is set to the same limits (in mg/dL) as for calcium.

CTC grades for the derived absolute WBC differential counts (neutrophils, lymphocytes) and corrected calcium will be assigned as described above for grading

5.3.3 Laboratory Notable Values

The liver function parameters of interest in this study are total bilirubin (TBIL), ALT, AST and alkaline phosphatase (ALP). The following notable criteria as per the most up-to-date version of the Novartis Liver Toxicity guidelines will be summarized:

- ALT > 3x ULN, 5x ULN, 10x ULN, 20 xULN
- AST > 3x ULN, 5x ULN, 10x ULN, 20 xULN
- ALT or AST > 3x ULN, 5x ULN, 8x ULN, 10x ULN, 20x ULN
- TBIL > 2x ULN, 3x ULN

For the following combined categories, the assessments need not be concurrent, i.e. participants are counted based on their most extreme value for each parameter (highest in the case of ALT, AST and TBIL; lowest in the case of ALP). Further medical review may be conducted to assess potential confounding factors such as liver metastases, liver function at baseline, etc.

- If ALT and AST ≤ ULN at baseline
  - ALT or AST > 3x ULN & TBL > 2x ULN
  - ALT or AST > 3x ULN & TBIL > 2x ULN & ALP ≥ 2x ULN
  - ALT or AST > 3x ULN & TBIL > 2x ULN & ALP < 2x ULN
- If ALT or AST > ULN at baseline
  - Elevated ALT or AST* & TBIL (>2x Baseline and 2x ULN)
  - Elevated ALT or AST* & TBIL (>2x Baseline and 2x ULN) & ALP ≥2x ULN
  - Elevated ALT or AST* & TBIL (>2x Baseline and 2x ULN) & ALP <2x ULN

* Elevated AST or ALT defined as: >3x ULN if ≤ ULN at baseline, or (>3x Baseline or 8x ULN) if > ULN at baseline

5.4 Other safety data

5.4.1 ECG notable values

For ECG data, the following notable values will be assessed:

- QT, QTcF
  - New value of > 450 and ≤ 480 ms
  - New value of > 480 and ≤ 500 ms
  - New value of > 500 ms
  - Increase from Baseline of > 30 ms to ≤ 60ms
  - Increase from Baseline of > 60 ms
- HR (for adults in Group 1)
  - Increase from baseline >25% and to a value > 100 bpm
  - Decrease from baseline >25% and to a value < 50 bpm
- HR (for pediatrics in Groups 2, 4 and 5)

Age category	**Notable value[a]**
2 - 5 years	- New[b] HR < 75 - New[b] HR > 125 bpm
6 - 11 years	- New[b] HR < 60 - New[b] HR > 100 bpm
12 - 17 years	- New[b] HR < 50 - New[b] HR > 100 bpm
≥ 18 years	- Increase from baseline >25% and to a value > 100 bpm - Decrease from baseline >25% and to a value < 50 bpm

[a] Notable values are based on age at the time of the assessment. For instance, if a participant is 11 years at screening/enrollment and is 12 years at Week 32, then notable values for age category "6 - 11 years" should be used for post-baseline ECG assessments at Weeks 4, 8, 12, 16 and 24. For ECG assessments at Week 36, 40, ..., notable values for age category "12 - 17 years" should be used.

[b] New means the value at baseline is not meeting the criterion.

- PR
  - Increase from baseline >25% and to a value > 200 ms
  - New value of > 200 ms
- QRS
  - Increase from baseline >25% and to a value > 120 ms
  - New value of QRS > 120 ms

5.4.2 Vital signs notable values

For analysis of vital signs the clinically notable vital sign criteria for adults (Group 1) are provided in Table 5-3 below.

Table 5-3 Clinically notable changes in vital signs for adults (Group 1)

Vital sign (unit)	Clinically notable criteria	
	above normal value	**below normal value**
Weight (kg)	increase > 10% from Baseline	decrease > 10% from Baseline
Systolic blood pressure (mmHg)	>=180 with increase from baseline of >=20	<=90 with decrease from baseline of >=20
Diastolic blood pressure (mmHg)	>=105 with increase from baseline of >=15	<=50 with decrease from baseline of >=15
Pulse rate (bpm)	>=100 with increase from baseline of >25%	<=50 with decrease from baseline of > 25%
Body temperature	>= 39.1	≤ 35.0°C

Clinically notable vital sign criteria for pediatric participants <18 years of age are provided Table 5-4 below.

Table 5-4 Clinically notable changes in vital signs for pediatrics (Groups 2, 4 & 5)

Vital sign		Participant age at visit		
		< 18 years		**≥ 18 years**
Systolic blood pressure [mmHg]	High	≥ 95th percentile of the age and height group[1]		≥ 180 with increase from updated baseline[5] of ≥20 mmHg
	Low	≤ 5th percentile of the age and height group[1]		≤ 90 with decrease from updated baseline[5] of ≥20 mmHg
Diastolic blood pressure [mmHg]	High	≥ 95th percentile of the age and height group[1]		≥ 105 with increase from updated baseline[5] of ≥15 mmHg
	Low	≤ 5th percentile of the age and height group[1]		≤ 50 with decrease from updated baseline[5] of ≥15 mmHg
Oral body temperature [°C]	High	≥ 38.4°C		≥ 39.1°C
	Low	≤ 35.0°C		
Pulse rate [bpm] [2]	High	2-3 years	> 128	≥120 with increase from updated baseline[5] of ≥15 bpm
		3-4 years	> 123	
		4-6 years	> 117	
		6-8 years	> 111	
		8-12 years	> 103	
		12-15 years	> 96	
		≥ 15 years	> 92	

Vital sign		Participant age at visit		
		< 18 years		≥ 18 years
	Low	2-3 years	< 92	≤50 with decrease from updated baseline[5] of ≥15 bpm
		3-4 years	< 86	
		4-6 years	< 81	
		6-8 years	< 74	
		8-12 years	< 67	
		12-15 years	< 62	
		≥ 15 years	< 58	
Weight	High	increase from baseline[3] of ≥ 2 BMI-for-age percentile categories[4]		Weight increase from updated baseline[5] of ≥ 10%
	Low	decrease from baseline[3] of ≥ 2 BMI-for-age percentile categories[4]		Weight decrease from updated baseline[5] of ≥ 10%

bpm=beats per minute;

[1] Blood pressure percentiles are calculated for each blood BP record using the method described in Appendix B of the following reference: The Fourth Report on Diagnosis, Evaluation and Treatment of High Blood Pressure in Children and Adolescents. Pediatrics 2004; 114; 555.

[2] Fleming S, Thompson M, Stevens R, et al. Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies. Lancet 2011; published online March 15. DOI:10.1016/S0140-6736(10)62226-X.

[3] Baseline BMI-for-age weight status categories are underweight (less than the 5th percentile), healthy weight (5th percentile to less than the 85th percentile), overweight (85th to less than the 95th percentile) and obese (equal to or greater than the 95th percentile);

[4] BMI-for-age percentiles categories (P3, P5, P10, P25, P50, P75, P85, P90, P95, P97) are obtained from the WHO Growth Charts (http://www.who.int/childgrowth/en/);

[5] Updated baseline is the last value collected before the 18th birthday.

5.4.3 Groupings of prior and concomitant medications

Medications for coagulopathy will be defined in the TFLs based on clinical review.

Pain medications are reported on the prior and concomitant medications CRF in the category "Pain relief medications". Pain medications will be classified according to ATC code into the following categories:

Non-opioids (including salicylates/paracetamol and NSAID): e.g., paracetamol (N02BE), nonsteroidal anti-inflammatory drugs (NSAIDs, M01A: ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS, NON-STEROIDS), metamizole (N02BB), M02: TOPICAL PRODUCTS FOR JOINT AND MUSCULAR PAIN, M03: MUSCLE RELAXANTS, V10B: PAIN PALLIATION, M09AB: chymopapain, A03D: ANTISPASMODICS IN COMBINATION WITH ANALGESICS.

Opioids: Opioids are defined as medications with ATC code in N02A. If there are sufficient data, opioids may be further categorized into weak and strong opioids.

Weak opioids: oral or rectal opioid formulations with a morphine conversion factor of 0.3 or less include (see e.g., Wertli et al 2017): codeine and combinations (N02AJ), N02AX (tilidine, tramadol, tapentadol).

Strong opioids: all opioids not defined as weak (morphine, oxycodone, hydromorphone, methadone, ...). Opioids are defined as medications with ATC code in N02A: Opioids (excluding above list of weak opioids).

If a participant concomitantly receives several pain medications in different categories, the highest category based on the individual medications will be applied.

After medical review/coding these lists of ATC codes for pain medications may be extended. Further sub-categories (e.g. based on defined daily dose for opioids) may be identified.

5.5 Patient reported outcomes

5.5.1 Visit windows

Time windows will be defined for descriptive summary of PRO data by visit and longitudinal data analysis. If more than one assessment is available in the same time window, the assessment closest to the planned date will be considered. If two assessments are obtained with the same time difference compared to the scheduled visit day, the assessment obtained prior to visit will be considered. If two assessments (scheduled and unscheduled) are on the same date, then the scheduled assessment will be considered. Data obtained at the end of treatment will be classified as other assessment in the corresponding time window. Note that only data collected under treatment (i.e. while the patient is treated) will be included in the analysis. The end of treatment assessment will be included if collected within 30 days of the last dose intake.

For instance, Table 5-5 shows time windows for PGI-S assessments. The mid-point between the planned study day for two consecutive visits is taken to define time windows. For PRO assessments with a visit schedule different to the PGI-S, time windows based on mid-points will be defined in a similar manner.

Table 5-5 Time windows for PRO: PGI-S

Time Window	Planned Visit Timing	Time Window Definition
On treatment		
Baseline (Week 1 Day 1)	*On or before Study Day 1**	*≤ Study Day 1*
Week 2 Day 1	*Study Day 8*	*Study Days 2 - 18*
Week 4	*Study Day 28*	*Study Days 19 - 41*
Week 8	*Study Day 56*	*Study Days 42 - 83*
Week 16	*Study Day 112*	*Study Days 84 - 135*
Week 20	*Study Day 140*	*Study Days 126 - 153*
Week 24	*Study Day 168*	*Study Days 154 - 181*
Week 28	*Study Day 196*	*Study Days 182 - 209*
Week 32	*Study Day 224*	*Study Days 210 - 251*
Week 40	*Study Day 280*	*Study Days 252 - 307*
Week 48	*Study Day 336*	*Study Days 308 - 419*
Week 72	*Study Day 504*	*Study Days 420 - 587*
Every 24 weeks thereafter		
*Week y=72+24*k*	*Study Day (72+24*k)*7*	*Study Days*
(with k = 1, 2, …)		*(72+24*k)*7 - 84 to*

		*(72+24*k)*7 + 83*
End of treatment		
End of treatment	*N.A.*	*Data collected under EOT visit, if no data were collected at the EOT visit last available data obtained before EOT*
** Study Day 1 = randomization date*		

5.5.2 Derivation of total scores

5.5.2.1 Patient Global impression of Symptom Severity (PGI-S)

The PGI-S is a patient-reported instrument that measures the overall severity of symptoms experienced on a 5-point scale ranging from "no symptoms" to "very severe".

The PGI-S corresponds to a single item. Derivation of total scores is not applicable.

5.5.2.2 Evaluation of Pain intensity, location and type (collected with PRO diary)

The Brief Pain Inventory (BPI) is a multi-item assessment of pain; the following items from this questionnaire will be used:

- worst pain intensity in the past 24 hours, assessed on an 11-point response scale ranging from 0 (no pain) to 10 (pain as bad as you can imagine)
- interference with general activity in the past 24 hours, assessed on an 11-point response scale ranging from 0 (does not interfere) to 10 (completely interferes)
- location of pain on a modified BPI diagram
- type of pain experienced (e.g., aching, throbbing)

The items will be presented separately. A total score will not be derived.

5.5.2.3 Dyspnea items (collected with PRO diary)

Two items on Dyspnea Severity in the PRO diary (included based on cognitive debriefing interviews) assess the degree to which respondents felt short of breath while completing various physical activities:

- shortness of breath when walking up 10 stairs
- shortness of breath when talking while walking

Each of the items includes 5 response options: no shortness of breath (0), mildly short of breath (1), moderately short of breath (2), severely short of breath (3), and I did not do this.

The two items will be presented separately. A total score will not be derived.

5.5.2.4 PROMIS Profile domains

The PROMIS Profiles are a group of PROMIS short forms measuring different domains of HRQoL from the Patient-Reported Outcomes Measurement Information System (PROMIS) system (Ader and Deborah, 2007). The age-specific tools and PROMIS domains and items are further described in the protocol, Section 2.11 and Table 5-6.

Table 5-6 PROMIS measures by age group (cf. Protocol Table 8-8)

Concept of interest	Adult Self-report (≥18 years old)	Child Self-report (≥12-17 years old)	Child Report with Parent/Caregiver Assistance (3 - 11 years old)	Parent/Caregiver Reporter (<12 years old)
	PRO measure			
Physical function	PROMIS-29+2 Profile v2.1	PROMIS Pediatric-25 Profile v2.0	NA	PROMIS-Parent-Proxy-25 Profile v2.0
Fatigue				
Ability to participate in social/peer relationships[1]				
Pain interference				
Pain intensity				
Anxiety				
Depression				
Cognitive function		NA	NA	NA
Sleep disturbance[2]		PROMIS Pediatric Short Form v1.0 - Sleep Disturbance 4a	NA	PROMIS Parent Proxy Short Form v1.0 -Sleep Disturbance 4a

[1] Domain "ability to participate in social roles and activities" is assessed for adults, while "peer relationships" is assessed for children/parent-proxy
[2] Sleep disturbance is assessed in the PROMIS-29 Profile v2.1 using the PROMIS Sleep Disturbance Short Form 4a. Sleep disturbance is not assessed on the PROMIS pediatric or parent profile version, and thus a separate short form assessment is needed.

Scoring of PROMIS domains will be done according to the PROMIS Scoring Manuals (for PROMIS Adult Profile Instruments, and for PROMIS Pediatric and Parent Proxy Profile Instruments). Scoring of the PROMIS Sleep Disturbance instrument will be done according to the PROMIS Scoring manual for Sleep Disturbance instruments.

Each of the PROMIS Profiles includes an additional pain intensity 0-10 numeric rating scale (NRS). This item will be presented separately, and no total score will be derived.

PROMIS instruments are scored using item-level calibrations. Each question usually has five response options ranging in value from one to five. To find the total raw score for a short form with all questions answered, sum the values of the response to each question. For example, for the adult profile 8-item form, the lowest possible raw score is 8; the highest possible raw score is 40 (see all short form scoring tables in Appendix 1 of the PROMIS manuals). All questions must be answered in order to produce a valid score using the scoring tables.

The applicable score conversion tables are provided in Appendix 1 of the manuals, and they can be used to translate the total raw score into a T-score for each participant. The T-score rescales the raw score into a standardized T-score with a mean of 50 and a standard deviation (SD) of 10. Therefore, a person with a T-score of 40 is one SD below the mean.

For example for the PROMIS Pediatric – 25 Pediatric Profile 2.0 instrument, an Anxiety raw score of 10 converts to a T-score of 54.8 with a standard error (SE) of 5 (see Table 5-7). Thus, the 95% confidence interval around the observed score ranges from 45 to 64.6 (T-score + (1.96*SE) or 54.8 + (1.96*5).

For example, for the PROMIS – 29 Adult Profile 2.1 instrument, an Anxiety raw score of 10 converts to a T-score of 59.5 with a standard error (SE) of 2.6 (see scoring table for the 4a short form v2.1 in Appendix 1 of the manual). Thus, the 95% confidence interval around the observed score ranges from 54.4 to 64.6 (T-score + (1.96*SE) or 59.5 + (1.96* 2.6).

For most PROMIS instruments, a score of 50 is the average for the United States general population with a standard deviation of 10 because calibration testing was performed on a large sample of the general population.

A *higher PROMIS T-score represents more of the concept being measured*. For negatively-worded concepts like Anxiety, a T-score of 60 is one SD worse than average. By comparison, an Anxiety T-score of 40 is one SD better than average. However, for positively-worded concepts like Physical Function-Mobility, a T-score of 60 is one SD better than average while a T-score of 40 is one SD worse than average.

Table 5-7 Illustration of calculating T-scores from raw scores based on Appendix 1 of the PROMIS scoring manual for Pediatric – 25 v2.0 and Parent Proxy-25 v2.0 instruments

PROMIS Pediatric – 25 v2.0

Anxiety 4b *Short Form Conversion Table*		
Raw Score	**T-score**	**SE***
4	35.6	6.4
5	40.9	5.6
6	44.1	5.4
7	47.2	5.2
8	49.9	5.1
9	52.4	5
10	54.8	5
11	57.2	5
12	59.5	5
13	61.8	5
14	64	5.1
15	66.3	5.1
16	68.7	5.1
17	71.1	5.1
18	73.7	5.2
19	76.3	5.1
20	79.5	5.1
*SE = Standard Error		

PROMIS Parent Proxy – 25 v2.0

Anxiety 4a *Short Form Conversion Table*		
Raw Summed Score	T-score	SE*
4	36.3	6
5	42.3	4.9
6	46	4.5
7	49.1	4.2
8	52	4
9	54.8	3.9
10	57.4	3.9
11	60	4
12	62.7	4
13	65.4	4.1
14	68	4.1
15	70.6	4.1
16	73.1	4
17	75.6	3.9
18	78.2	3.9
19	80.8	3.8
20	83.6	3.6
*SE = Standard Error on T-score metric		

Source: PROMIS Scoring Manual for Pediatric and Parent Proxy Profile Instruments.

5.5.3 Derivation of weekly worst pain intensity, weekly pain interference, dyspnea items based on PRO diary

The PRO diary will be used to collect daily information on pain as follows:

- Worst pain intensity in the last 24 hours
- Pain interference with general activity during the last 24 hours
- primary location of pain and associated type of pain

Worst pain intensity, and associated type of pain will also be asked for a secondary location of pain.

The diary also includes two dyspnea items.

For the assessment of pain and dyspnea, participants (or their caregivers) will be asked to complete these items on the PRO Diary on the ePRO devices every evening starting in the screening period (starting 7 days prior the first dose of study treatment) through week 16, and again starting at week 20 through week 24.

Worst pain intensity, pain interference and dyspnea items will be averaged weekly over a 7-day period if a participant has completed the questionnaire for at least 4 days in the 7-day period

prior to/at the time points (see Table 5-9). The weekly mean will be calculated based on the available assessments.

Table 5-8 Weekly worst pain intensity, pain interference, and dyspnea items windows

Time point	Time Interval for weekly PRO average
Baseline	Day -7 to Day -1
Week 1	Day 1 to day 7
Week *k* (with *k*= 2, …,16, 20,…, 24)	Day 1+(*k*-1)*7 to day *k**7

6 Reference

Adams DM, Trenor CC, Hammill AM, et al (2016) Efficacy and Safety of Sirolimus in the Treatment of Complicated Vascular Anomalies. Pediatrics p. e20153257.

Ader, Deborah N. (2007) Developing the Patient-Reported Outcomes Measurement Information System (PROMIS);45(5):S1-S2

Baumgartner RN, Roche AF, Himes (1986). Incremental growth tables: supplementary to previously published charts. American Journal of Clinical Nutrition, 43, 711-722.

Boechat MI, Lee DC (2007) Graphic Representation of Skeletal Maturity Determinations, American Journal of Roentgenology. 189: 873-874.

Bretz F, Maurer W, Brannath W, et al (2009) A graphical approach to sequentially rejective multiple test procedures. Stat Med p. 586-604.

Canaud C, Gutiérrez JCL, Irvine A, et al (2021) EPIK-P1: Retrospective chart review study of patients (pts) with PIK3CA-related Overgrowth Spectrum (PROS) who have received alpelisib (ALP) as part of a compassionate use programme [abstract]. Annals of Oncology; 32 Suppl5:S1297.

Carel JC, Lahlou N, Roger M, Chaussain JL (2004). Precocious puberty and statural growth. Hum Reprod Update; 10(2):135-47.

Cella D, Choi SW, Condon DM, et al (2019) PROMIS® Adult Health Profiles: Efficient Short-Form Measures of Seven Health Domains. Value Health p. 537-544.

Fenichel, P (2012). Delayed puberty. Endocrine Development, 22, 138–159.

Greulich, WW and Pyle SI (1959) Radiographic Atlas of Skeletal Development of the Hand and Wrist; 8(4):513

Hernán MA and Robins JM (2020) Causal Inference: What If. 1st edition. Boca Raton: Chapman & Hall/CRC, chapter 12 and 13. https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/

Hojreh A, Gamper J, Schmook MT, et al (2018) Hand MRI and the Greulich-Pyle atlas in skeletal age estimation in adolescents. Skeletal Radiol. 47(7):963-71.

Marshall, WA, & Tanner, JM (1969). Variations in pattern of pubertal changes in girls. Archives of disease in childhood, 44(235), 291–303.

Marshall, WA, & Tanner, JM (1970). Variations in the pattern of pubertal changes in boys. Archives of disease in childhood, 45(239), 13–23.

Maurer W, Glimm E, Bretz F (2011) Multiple and Repeated Testing of Primary, Coprimary, and Secondary Hypotheses, Statistics in Biopharmaceutical Research, 3:2, p. 336-352.

Midyett LK, Moore WV, Jacobson JD (2003). Are pubertal changes in girls before age 8 benign? Pediatrics; 111(1):47–51

Palmert MR, Dunkel L (2012). Clinical practice. Delayed puberty. N Engl J Med;366:443–53

Parker VER, Keppler-Noreuil KM, Faivre L, et al (2019) Safety and efficacy of low-dose sirolimus in the PIK3CA-related overgrowth spectrum. Genet. Med. p. 1189-1198.

Sasaki T, Ishibashi Y, Okamura Y, et al (2002) MRI evaluation of growth plate closure rate and pattern in the normal knee joint. J Knee Surg;15(2):72-6.

Venot Q, Blanc T, Rabia SH, et al (2018) Targeted therapy in patients with PIK3CA-related overgrowth syndrome. Nature p. 540-546.

Wertli, MM, Reich, O, Signorell, A et al (2017) Changes over time in prescription practices of pain medications in Switzerland between 2006 and 2013: an analysis of insurance claims. BMC Health Serv Res 17, 167.

Wong DL, Baker CM (1988) Pain in children: comparison of assessment scales. Okla Nurse p. 8.

World Health Organization. Training Course on Child Growth Assessment. Geneva, WHO, 2008. https://apps.who.int/iris/bitstream/handle/10665/43601/9789241595070_C_eng.pdf