

Protocol A5481092

Phase 1/2 Study to Evaluate Palbociclib (IBRANCE®) in Combination with Irinotecan and Temozolomide or in Combination with Topotecan and Cyclophosphamide in Pediatric Patients with Recurrent or Refractory Solid Tumors

Statistical Analysis Plan (SAP)

Version: **5.0**

Date: **25-Aug-2023**

This document contains confidential information belonging to Pfizer Inc. Except as may be otherwise agreed to in writing, by accepting or reviewing these materials, you agree to hold such information in confidence and not to disclose it to others (except where required by applicable law), nor to use it for unauthorized purposes. In the event of actual or suspected breach of this obligation, Pfizer Inc. should be promptly notified.

TABLE OF CONTENT

LIST OF TABLES 6
LIST OF FIGURES 6
APPENDICES 7
1. VERSION HISTORY 8
2. INTRODUCTION 10
2.1. Study Objectives 10
2.1.1. Primary Objectives 10
2.1.2. Secondary Objectives 11
2.1.3. Exploratory Objectives 12
2.2. Study Design 13
2.2.1. Phase 1 Portion - Combination of Palbociclib With IRN and TMZ 13
2.2.2. Phase 1 Portion - Combination of Palbociclib With TOPO and CTX 14
2.2.3. Randomized Phase 2 Portion 16
3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS 16
3.1. Primary Endpoints 16
3.2. Secondary Endpoints 17
3.2.1. Safety Endpoints 17
3.2.2. Efficacy Endpoints 18
3.2.3. Pharmacokinetic Parameters of Palbociclib, IRN, TMZ, TOPO, and CTX: 19
3.2.4. Quality of Life (QoL) Endpoints (Phase 2 Portion Only) 20
3.3. Exploratory Endpoints 21
3.4. Baseline Variables 21
3.4.1. Investigational Product, Study Drug, Study Treatment and Baseline Definitions 21
3.4.2. Baseline Characteristics 24
3.5. Safety Endpoints 24
3.5.1. Adverse Events 24
3.5.1.1. Treatment-Emergent Adverse Events (TEAEs) 24
3.5.2. Laboratory Data 24
4. ANALYSIS SETS 25

4.1. Full Analysis Set 25
4.2. Safety Analysis Set....25
4.3. Other Analysis Set....26
4.3.1. Per-Protocol Analysis Set (DLT Evaluable Set) 26
4.4. PK Analysis Set....27
4.5. Biomarker Analysis Set....27
5. GENERAL METHODOLOGY AND CONVENTIONS....27
5.1. Hypotheses and Decision Rules 27
5.1.1. Hypotheses and Sample Size Determination 27
5.1.2. Statistical Decision Rules 29
5.2. General Methods 32
5.2.1. Analyses for Categorical Data 33
5.2.2. Analyses for Continuous Data 34
5.2.3. Analyses for Time to Event Data....34
5.2.4. Definition of Study Day....34
5.2.5. Definition of On-Treatment Period 35
5.2.6. Definition of Start of New Anticancer Therapy 35
5.2.7. Standard Derivation and Reporting Conventions 35
5.2.8. Unscheduled Visits 36
5.2.9. Adequate Baseline Tumor Assessment 36
5.2.10. Adequate Post-Baseline Tumor Assessment 36
5.3. Methods to Manage Missing Data 36
5.3.1. Missing or Partial Dates....36
5.3.1.1. Adverse Events....37
5.3.1.2. Exposure....37
5.3.1.3. Date of Last Contact....38
5.3.1.4. Death Date....38
5.3.1.5. Tumor Assessments....38
5.3.1.6. Missing Data in Time-to-Event Endpoints....39
5.3.2. Missing QTc Data....39
5.3.3. Pharmacokinetic Concentrations 40
5.3.4. Pharmacokinetic Parameters....40
5.3.5. Missing PRO Data 40

5.3.6. Missing/Duplicate Biomarker Data 41
6. ANALYSES AND SUMMARIES 41
6.1. Primary Endpoints 41
6.1.1. First Cycle DLT (Phase 1 Portion) 41
6.1.1.1. Primary analysis 41
6.1.2. Adverse Events (Phase 1 Portion and Tumor-Specific Cohorts) 41
6.1.2.1. Primary analysis 41
6.1.3. Objective Response as Assessed by the Investigator for the Dose Expansion Cohort (Phase 1 Portion) and Tumor-Specific Cohort(s) 42
6.1.3.1. Primary analysis 42
6.1.4. EFS Based on Investigator Assessment (Phase 2 Portion) 43
6.1.4.1. Primary analysis 43
6.1.4.2. Sensitivity analyses 45
6.2. Secondary Endpoints 46
6.2.1. Safety Endpoints 46
6.2.2. Efficacy Endpoints 46
6.2.2.1. EFS assessed by the BIRC 47
6.2.2.2. OR 47
6.2.2.3. DOR 48
6.2.2.4. PFS 49
6.2.2.5. OS 50
6.2.2.6. PET-CT/MRI response after Cycle 4 (Phase 2 portion only) 51
6.2.3. PK Endpoints 51
6.2.3.1. PK Analysis of Palbociclib, IRN, TMZ, TOPO and CTX 51
6.2.3.2. Population PK Analysis or PK/PD Modeling 52
6.3. Other Endpoints 52
6.3.1. PRO Endpoints (Phase 2 portion) 52
6.3.1.1. PRO questionnaires 52
6.3.1.2. PRO scoring 53
6.3.1.3. PRO questionnaire completion status 53
6.3.1.4. Descriptive summary 53
6.3.2. Days of hospitalization (Phase 2 portion) 53

6.3.3. Exploratory Biomarker Endpoints 53
6.3.4. Palatability of Palbociclib Oral Formulation 54
6.3.5. Bone Metabolism and Growth 54
6.4. Subset Analyses 54
6.5. Baseline and Other Summaries and Analyses 55
6.5.1. Baseline Summaries 55
6.5.2. Study Conduct and Patient Disposition 56
6.5.2.1. Patient Evaluation Groups 56
6.5.2.2. Patient Disposition 57
6.5.3. Protocol Deviations 57
6.5.4. Impact of COVID-19 57
6.5.5. Study Treatment Exposure and Compliance 57
6.5.5.1. Exposure to study drug 58
6.5.5.2. Treatment Modifications 58
6.5.6. Concomitant Medications and Non-Drug Treatments 59
6.5.7. Follow-Up Cancer Therapies 59
6.6. Safety Summaries and Analyses 59
6.6.1. Adverse Events 60
6.6.1.1. All Adverse Events and Serious Adverse Events 60
6.6.1.2. Adverse Events Leading to Dose Modification 61
6.6.1.3. Adverse Events Leading to Permanent Discontinuation of Study Treatment 62
6.6.1.4. Adverse Events Leading to Death 62
6.6.2. Laboratory Data 62
6.6.3. Vital Signs and Physical Examination 63
6.6.4. ECG 63
7. INTERIM ANALYSES 64
7.1. Introduction 64
7.2. Interim Analyses and Summaries 65
8. REFERENCES 66
9. APPENDICES 68

LIST OF TABLES

Table 1. Summary of Major Changes in SAP Amendments 8
Table 2. Dose Levels for Palbociclib, IRN and TMZ....14
Table 3. Dose Levels for Palbociclib, TOPO and CTX....15
Table 4. Decision Rules in Rolling 6 Design for Combination of Palbociclib with IRN and TMZ 30
Table 5. Decision Rules in Rolling 6 Design for Combination of Palbociclib with TOPO and CTX 31
Table 6. Modified Simon's Optimal 2-Stage Design Decision Rules for Tumor-Specific Cohorts 32
Table 7. Study Summaries and Tabulations for Each Treatment Combination - Phase 1 Portion and Tumor-Specific Cohort(s) 33
Table 8. Statistical Analyses by Analysis Set - Phase 2 Portion 34
Table 9. EFS Censoring Reasons and Hierarchy....44
Table 10. OS Censoring Reasons and Hierarchy....50

LIST OF FIGURES

Figure 1. Study Schema - Phase 1 Palbociclib in Combination With IRN and TMZ....13
Figure 2. Study Schema - Phase 2 Palbociclib in Combination With IRN and TMZ....16

APPENDICES

Appendix 1. RECIST (Response Evaluation Criteria in Solid Tumors) version 1.1 Guidelines 68
Appendix 2. Modified RANO Criteria for Patients with CNS Malignancies 73
Appendix 3. Response Criteria for Neuroblastoma 77
Appendix 4. Data Derivation Details 84
Appendix 5. Study Treatment Modification and Compliance Derivation Details 85
Appendix 6. Palbociclib Oral Capsule Dosing Table 89
Appendix 7. Volume of Palbociclib Oral Solution to be Administered by Dose and BSA 90
Appendix 8. Dosing Nomogram for Oral Temozolomide 91
Appendix 9. Palatability Questionnaire 92
Appendix 10. List of Abbreviations 94

1. VERSION HISTORY

This Statistical Analysis Plan (SAP) Amendment for study A5481092 is based on the Protocol Amendment 7 dated 31May2023 and driven by the blinded review of the data.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Version Date	Summary of Changes
1.0	31-Jan-2019	Not Applicable
2.0	02-Nov-2020	As per protocol amendment 3, updates were made to the title and the relevant sections of the SAP to incorporate analyses associated with the additional combination of palbociclib with topotecan and cyclophosphamide chemotherapy backbone; added to the relevant sections to include analyses associated with the additional Ewing sarcoma tumor-specific cohort for palbociclib in combination with irinotecan and temozolomide. Added updates as per protocol amendment 2 including additional dose levels to the dose escalation part for combination of palbociclib with irinotecan and temozolomide, and additional tumor response criteria for neuroblastoma. Added definition of the primary analysis cutoff date. Included minor editorial and consistency changes throughout the document.
3.0	07-Sep-2021	As per protocol amendment 4, updates were made to the study title and the relevant sections of the SAP to incorporate analyses associated with the inclusion of the randomized Phase 2 portion of the study. Included minor editorial and consistency changes throughout the document.
4.0	08-Jul-2022	As per protocol amendment 5, updates were made to the study title and the relevant sections associated with the inclusion of event-free survival assessed by an independent review committee as a secondary endpoint in the Phase 2 portion of the study. Added sensitivity analysis of EFS in patients with molecularly confirmed EWS. Updated objective response status of NE in tumor assessments after concurrent surgery for local control and potential censoring.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Version Date	Summary of Changes
		Included objective response status for bone marrow assessment to confirm CR in the Phase 2 portion. Updated definition of hospitalization and added admittance reasons to the analysis of hospitalization in the Phase 2 portion. Additional editorial and consistency changes throughout the document.
5.0	24-Aug-2023	Updates were made to the following sections: Section 2 - Added definition of the data cutoff date for the tumor-specific cohort based on adequate evaluation of tumor response. Added definition of the final close-out analysis. Section 3.4.2 Baseline Characteristics and Section 6.4 Subset Analysis - Added clarity to stratification factor with respect to time of current disease recurrence since completion of first line treatment. Section 4.3.1 Per-Protocol Analysis Set – Updated wording “Evaluable for DLT” to “DLT Evaluable Set”. Section 5.2.9 Adequate Baseline Tumor Assessment – Updated definition of adequate baseline assessment. Section 6.1.1.1 Primary Analysis (Phase 1 portion) – Added analysis of first cycle DLTs by MedDRA PT including clusters of PTs. Section 6.3.5 Bone Metabolism and Growth – Added additional analysis of bone age vs. calendar age. Section 6.5.1 Baseline Summaries – Removed “BMI” and “intent” (neoadjuvant, adjuvant) for prior anticancer drug therapy; added “Site of recurrence” and “Measurable disease” for Phase 2 portion due to clinical relevance. Section 6.5.4 Impact of COVID-19 – Added summary of patient discontinuations related to COVID-19. Section 6.6.1.1 All AEs and SAEs – Included clusters of PTs in the relevant summaries of TEAEs and treatment-emergent SAEs.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Version Date	Summary of Changes
		Additional editorial and consistency changes throughout the document.

2. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in Study A5481092, a Phase 1/2 study to evaluate palbociclib in combination with irinotecan (IRN) and temozolomide (TMZ) or in combination with topotecan (TOPO) and cyclophosphamide (CTX) in pediatric patients with recurrent or refractory solid tumors. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment. Any deviations from this analysis plan will be described in the Clinical Study Report.

Statistical analyses will be performed using cleaned eCRF data as well as non-CRF data (eg, pharmacokinetic [PK], biomarker data, central testing data for molecular diagnosis, tumor assessment and pathology data from the Blinded Independent Review Committee [BIRC]). The primary analysis will include all data up to a cutoff date which is determined by the number of events required for the primary analysis of event-free survival (EFS) based on investigator assessment in the Phase 2 portion of the study. Corresponding data snapshots will be taken to allow cleaning of data up to and including the cutoff date. Due to data cleaning activities, the final number of EFS events might deviate from the planned number. The data cutoff date will not be adjusted retrospectively in this case.

The data cutoff date for the primary analysis of the tumor-specific cohort is determined by the minimum follow-up time required for adequate assessment of tumor response justified by clinical relevance, which corresponds to 6 months after the last patient receives the first dose of study treatment. This data cutoff also applies to tumor response evaluation in Stage 1 unless criteria for proceeding to Stage 2 have already been met.

The final close-out analysis will be performed after last patient last visit (LPLV).

2.1. Study Objectives

2.1.1. Primary Objectives

Phase 1 Portion - Dose Escalation Cohort

- To estimate the maximum tolerable dose (MTD) for the combination of palbociclib, TMZ and IRN in in children, adolescents, and young adults with recurrent or refractory solid tumors.

Phase 1 Portion - Dose Determination Part

- To determine the potential recommended Phase 2 dose (RP2D) for palbociclib in combination with TOPO and CTX in children, adolescents, and young adults with recurrent or refractory solid tumors.

Phase 1 Portion - Dose Expansion Cohorts and Tumor-Specific Cohort(s)

- To evaluate the safety, and confirm the RP2D for the combination of palbociclib, IRN and TMZ at the MTD in children, adolescents and young adults with recurrent or refractory solid tumors that are known to be sensitive to a CDK4/6 inhibitor.

- To evaluate the preliminary antitumor activity of palbociclib combined with IRN and TMZ in children, adolescents and young adults with recurrent or refractory solid tumors (Dose Expansion Cohort/Tumor-Specific Cohort excluding EWS).

- To evaluate the safety, and confirm the RP2D for the combination of palbociclib, TOPO and CTX in children, adolescents and young adults with recurrent or refractory solid tumors that are known to be sensitive to a CDK4/6 inhibitor.

- To evaluate the preliminary antitumor activity of palbociclib combined with TOPO and CTX in children, adolescents and young adults with recurrent or refractory tumor types.

Randomized Phase 2 Portion

- To compare the efficacy of palbociclib in combination with IRN and TMZ vs IRN and TMZ chemotherapy alone in the treatment of children, adolescents, and young adults with recurrent or refractory EWS.

2.1.2. Secondary Objectives

Phase 1 Portion and Tumor-Specific Cohorts

- To characterize the safety profile of palbociclib combined with IRN and TMZ in children, adolescents and young adults with recurrent or refractory solid tumors (Dose Escalation Part)

- To describe the PK of palbociclib, IRN and TMZ in children, adolescents, and young adults with recurrent or refractory solid tumors when given in combination. (Dose Escalation Part/Dose Expansion Cohort/Tumor-Specific Cohorts)

- To evaluate the preliminary antitumor activity of palbociclib combined with IRN and TMZ in children, adolescents and young adults with recurrent or refractory solid tumors. (Dose Escalation Part)

- To evaluate other efficacy parameters of palbociclib combined with IRN and TMZ in children, adolescents, and young adults with recurrent or refractory solid tumors. (Dose Expansion Cohort/Tumor-Specific Cohorts)

- To characterize the safety profile of palbociclib combined with TOPO and CTX in children, adolescents, and young adults with recurrent or refractory solid tumors. (Dose Determination Part)

- To describe the PK of palbociclib, TOPO and CTX in children, adolescents, and young adults with recurrent or refractory solid tumors when given in combination. (Dose Determination Part/Dose Expansion Cohort/Tumor-Specific Cohorts)

- To evaluate the preliminary antitumor activity of palbociclib combined with TOPO and CTX in children, adolescents, and young adults with recurrent or refractory solid tumors. (Dose Determination Part)

- To evaluate other efficacy parameters of palbociclib combined with TOPO and CTX in children, adolescents, and young adults with recurrent or refractory solid tumors. (Dose Expansion Cohort/Tumor-Specific Cohorts)

Randomized Phase 2 Portion

- To further compare the efficacy of palbociclib in combination with TMZ and IRN vs TMZ and IRN chemotherapy alone.

- To characterize the toxicity and safety of the combination of TMZ and IRN plus or minus palbociclib.

- To describe the PK of palbociclib, TMZ, and IRN in children, adolescents, and young adults with recurrent or refractory EWS when given in combination.

- To characterize the impact of the combination of palbociclib with TMZ and IRN treatment on the quality of life (QoL) of patients with refractory and recurrent EWS.

2.1.3. Exploratory Objectives

- To characterize candidate biomarkers of sensitivity or resistance in pre-treatment tumor tissue that may aid in the identification of patient subpopulations most likely to benefit from treatment.

- To explore pharmacodynamics (PD) modulation of thymidine kinase (TK)-1 activity in blood following palbociclib exposure.

- To evaluate circulating nucleic acid (CNA) biomarkers in blood.

- To evaluate preliminary information on the palatability of the oral solution formulation of palbociclib.

- To explore the effect of treatment with palbociclib and parameters of bone metabolism and growth, including linear growth, bone age, bone mineral density, physical signs of pubertal maturation in patients who are pre or peripubertal (<Tanner 4).

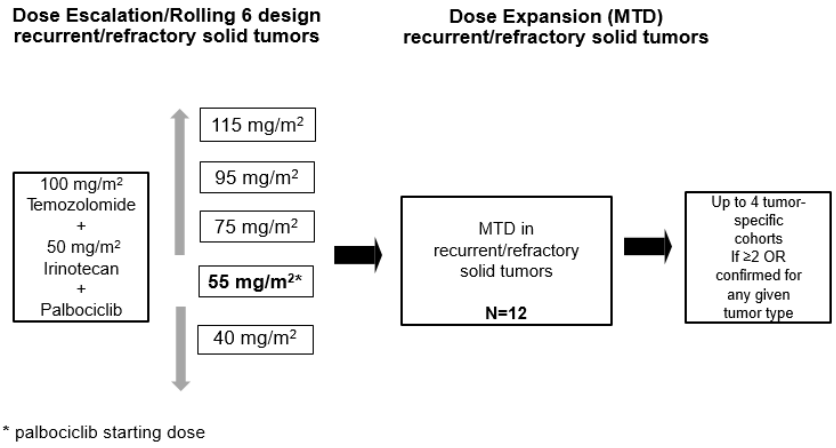
2.2. Study Design

This is a Phase 1/2 open-label, multicenter study to evaluate palbociclib in combination with IRN and TMZ or TOPO and CTX chemotherapy in children, adolescents and young adults with recurrent or refractory solid tumors. The study consists of a non-randomized Phase 1 portion in patients with recurrent or refractory solid tumors to determine the RP2D for the respective treatment combination, followed by a randomized Phase 2 portion to compare the efficacy of palbociclib in combination with TMZ and IRN vs TMZ and IRN chemotherapy alone in patients with recurrent or refractory EWS.

2.2.1. Phase 1 Portion - Combination of Palbociclib With IRN and TMZ

Evaluation of palbociclib in combination with IRN and TMZ consists of 3 parts (Figure 1): dose escalation part to estimate the MTD, dose expansion cohort to confirm the RP2D, and if applicable, tumor-specific cohort(s) other than EWS to further evaluate preliminary antitumor activity in the specific tumor type.

Figure 1. Study Schema - Phase 1 Palbociclib in Combination With IRN and TMZ



2.2.1.1 Dose Escalation Part

The dose levels of palbociclib in combination with IRN and TMZ are shown in Table 2 below.

The dose escalation part will follow a rolling 6 dose escalation design (Skolnik, 2008)[12] with the decision rules to guide dose escalation and to estimate the MTD.

The dose escalation part will enroll up to approximately 24 DLT-evaluable patients. The MTD is defined as the highest dose level at which <33% of a minimum of 6 evaluable patients experience a dose limiting toxicity (DLT) during Cycle 1 of treatment. If 6 or more patients have been treated at the highest dose level studied and <33% experience a DLT, then the MTD is not reached and the highest dose will proceed to dose expansion.

Table 2. Dose Levels for Palbociclib, IRN and TMZ

Dose Levels	Palbociclib Orally, once daily (mg/m^2) Day 1-14 in 21-Day Cycles	TMZ Orally (or IV if cannot swallow capsules), once daily (mg/m^2) Day 1-5 in 21-Day Cycles	IRN IV over 90 min, once daily, (mg/m^2) Day 1-5 in 21-Day Cycles
-1	40	100	50
1 (starting dose)	55	100	50
2	75	100	50
3	95	100	50
4	115	100	50

Abbreviations: IRN=irinotecan; IV=intravenous; TMZ=temozolomide.

2.2.1.2 Dose Expansion Cohort

Once the MTD has been estimated in the dose escalation part, a minimum of 12 patients with recurrent or refractory solid tumors will be enrolled and treated at the MTD in a dose expansion cohort in order to confirm the MTD as the RP2D.

The RP2D is the dose chosen for further investigation based on the results of dose escalation part and dose expansion cohort of the study. If the MTD proves to be clinically feasible for long-term administration in a reasonable number of patients, then this dose usually becomes the RP2D. Further experience with the MTD based on other safety, clinical activity, PK, and PD data may result in a RP2D dose lower than the MTD.

2.2.1.3 Tumor-Specific Cohort(s)

Once the RP2D is confirmed, preliminary antitumor activity of the combination of palbociclib with IRN and TMZ will be evaluated in the tumor-specific cohort(s) (maximum of 21 patients each) other than EWS (eg, neuroblastoma, rhabdomyosarcoma, rhabdoid tumor or medulloblastoma) which may be opened if 2 or more patients from the dose escalation part and/or dose expansion cohort show a confirmed objective response (OR) in a specific tumor type. As a separate Phase 2 portion will be conducted in patients with EWS, a tumor-specific cohort for EWS will not be opened for this combination.

2.2.2. Phase 1 Portion - Combination of Palbociclib With TOPO and CTX

Evaluation of palbociclib in combination with TOPO and CTX consists of 3 parts: dose determination part to determine the potential RP2D, dose expansion cohort to confirm the RP2D, and if applicable, tumor-specific cohort(s) to further evaluate preliminary antitumor activity in the specific tumor type.

2.2.2.1 Dose Determination Part

The dose levels of palbociclib in combination with TOPO and CTX are shown in Table 3 below.

Table 3. Dose Levels for Palbociclib, TOPO and CTX

Dose Levels	Palbociclib Orally, once daily (mg/m^2) Day 1-14 in 21-Day Cycles	TOPO (IV), over 30 min, once daily (mg/m^2) Day 1-5 in 21-Day Cycles	CTX (IV), over 30-60 min, once daily, (mg/m^2) Day 1-5 in 21-Day Cycles
-1	55	0.75	250
1 (starting dose)	75	0.75	250

Abbreviations: TOPO=topotecan; IV=intravenous; CTX=cyclophosphamide.

The dose determination part will follow a modified rolling 6 design with the decision rules to guide dose de-escalation and to determine the potential RP2D. The starting dose of palbociclib in combination with TOPO and CTX is the MTD determined in the combination of palbociclib with IRN and TMZ, where the MTD for this combination has been determined as 75 mg/m^2. This dose is considered the maximum administered dose (MAD) and no further dose escalation is planned.

The dose determination part will enroll approximately up to 12 DLT-evaluable patients. If <33% of a minimum of 6 evaluable patients experience a DLT at the MAD, the MAD will proceed to dose expansion to be confirmed as the RP2D. If ≥33% of evaluable patients experience a DLT, dose de-escalation to the next lower dose level will occur. If a minimum of 6 evaluable patients have been treated at the next lower dose level, the dose will proceed to dose expansion.

2.2.2.2 Dose Expansion Cohort

The dose expansion cohort will enroll a minimum of 12 patients with recurrent or refractory solid tumors to confirm the RP2D.

The study will aim to enroll 6 PK-evaluable patients per treatment combination in each of the following age groups: ≤6 years old, >6 years and <12 years old, and ≥12 years and <18 years old for each combination. The expansion cohorts may remain open for targeted enrollment by age group and by combination. However, given the relatively limited patient population of pediatric cancer patients in some of the age groups particularly the ≤6 and the >6 to <12 subgroups, the dose expansion cohorts may be closed if the available PK data is deemed sufficient per the Sponsor assessment.

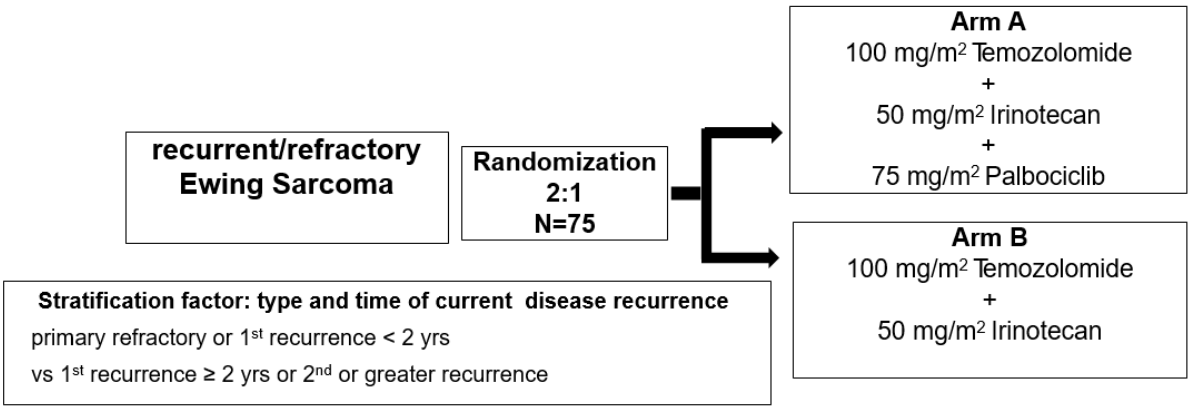
2.2.2.3 Tumor-Specific Cohort(s)

Once the RP2D is confirmed, combination of palbociclib with TOPO and CTX will be evaluated in the tumor-specific cohort(s) (maximum of 21 patients each), if 2 or more patients from the dose determination part and/or dose expansion cohort show a confirmed objective response in a specific tumor type (eg, EWS, neuroblastoma, rhabdomyosarcoma, rhabdoid tumor or medulloblastoma).

2.2.3. Randomized Phase 2 Portion

Once the RP2D is confirmed for the combination of palbociclib with IRN and TMZ in the Phase 1 portion of the study, the randomized Phase 2 portion of the study (Figure 2) will be initiated. Approximately 75 patients with recurrent or refractory EWS will be randomized in a 2:1 ratio to receive either palbociclib in combination with IRN and TMZ or IRN and TMZ chemotherapy alone. Randomization will be stratified using block randomization by type and time of current disease recurrence (primary refractory or 1st recurrence < 2 years vs. 1st recurrence ≥ 2 years since completion of first line treatment or 2nd or greater recurrence). Refractory disease is defined as disease progression during 1st line treatment or within 12 weeks of 1st line treatment completion. Disease progression at or after 12 weeks of completion of 1st line treatment is defined as recurrent disease.

Figure 2. Study Schema - Phase 2 Palbociclib in Combination With IRN and TMZ



3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoints

Phase 1 Portion - Dose Escalation Part/Dose Determination Part

- First cycle DLTs: DLTs during the first cycle of treatment which are attributable to at least one of the study drugs in the combination.

  The DLT evaluation period will be from Cycle 1 Day 1 until pre-dose Cycle 2 Day 1. Should there be a delay starting the subsequent cycle, the DLT evaluation period ends at pre-dose Cycle 2 Day 1.

Phase 1 Portion - Dose Expansion Cohorts and Tumor-Specific Cohorts

- Adverse events (AEs) as characterized by type, frequency, severity, timing, seriousness, and relationship to study treatment. Severity of AEs will be graded by the investigator according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.03 and coded using the latest version of the Medical Dictionary for Regulatory Activities (MedDRA) classification system.

- Laboratory test data including hematology and clinical chemistry as characterized by type, frequency, severity (as graded by NCI CTCAE v4.03) and timing.

- Electrocardiogram (ECG) parameters and vital signs (eg, blood pressure, pulse rate).

- OR as assessed by investigator using RECIST v1.1, modified Response Assessment in Neuro-Oncology (RANO) for CNS malignancies or International Neuroblastoma Response Criteria (INRC) for neuroblastoma.

  OR is defined as a best overall response (BOR) of confirmed complete response (CR) or partial response (PR) according to RECIST v1.1 or modified RANO. For neuroblastoma using INRC, OR is defined as a BOR of confirmed CR, PR, or minor response (MR). CR, PR, and MR must be confirmed by repeat assessments performed ≥ 4 weeks after the criteria for response are first met.

Randomized Phase 2 Portion

- EFS based on investigator assessment

  EFS is defined as the time from date of randomization to the date of first documentation of any of the following events based on investigator assessment:

  - Objective progressive disease per RECIST v1.1 (including progressive disease without achieving a response or recurrence following a response)
  - Diagnosis of secondary malignancy
  - Death due to any cause

Disease assessments will be performed as scheduled according to the study calendar to prevent the introduction of bias into the assessment of efficacy at screening, every 6 weeks (±7 days) from the date of randomization for the first 4 assessments, and every 12 weeks (±7 days) thereafter until first occurrence of disease progression as per RECIST v1.1 or discontinuation of patient from overall study participation (eg, death, patient's request, lost to follow-up), regardless of initiation of new anticancer therapy.

Patients who permanently discontinue study treatment for reasons other than disease progression as per RECIST v1.1 or secondary malignancy will continue to have tumor assessments every 6 weeks (±7 days) for the first 6 months since randomization, and then every 12 weeks (±7 days) thereafter until first occurrence of disease progression as per RECIST v1.1, diagnosis of secondary malignancy or discontinuation of patient from overall study participation (eg, death, patient's request, lost to follow-up), regardless of initiation of new anticancer therapy.

3.2. Secondary Endpoints

3.2.1. Safety Endpoints

Phase 1 Portion

For dose escalation/dose determination part:

- AEs as characterized by type, frequency, severity (as graded by NCI CTCAE v4.03), timing, seriousness, and relationship to study treatment.
- Laboratory test data as characterized by type, frequency, severity (as graded by NCI CTCAE v4.03) and timing.
- ECG parameters and vital signs.

Randomized Phase 2 Portion

- AEs as characterized by type, frequency, severity (as graded by NCI CTCAE v4.03), timing, seriousness, and relationship to study treatment.

3.2.2. Efficacy Endpoints

Phase 1 Portion and Tumor-Specific Cohort(s)

For dose escalation/dose determination part/dose expansion cohorts/tumor-specific cohorts:

- Duration of response (DOR) is defined for patients with confirmed OR (CR or PR) by investigator assessment, as the time from the date of first documentation of OR to the date of first documentation of progressive disease (PD) or death due to any cause.

For dose expansion cohorts and tumor-specific cohorts only:

- Progression-free survival (PFS) is defined as the time from the date of first dose of study treatment to the date of the first documentation of objective PD or death due to any cause, whichever occurs first.

  Post-baseline tumor assessments will be performed every 2 cycles (±7 days) for the first 8 cycles and then every 4 cycles (±7 days) from start of study treatment until first documented objective disease progression as per RECIST v1.1, or modified RANO for CNS malignancies or INRC for neuroblastoma, initiation of new anticancer therapy, or discontinuation of patient from overall study participation (eg, death, patient's request, lost to follow-up).

  Patients who permanently discontinue study medication for reasons other than objective disease progression as per RECIST v.1.1, modified RANO or INRC for neuroblastoma will continue to have tumor assessment performed during the follow-up visits every 6 weeks (±7 days) from last on-treatment tumor assessment for the first 6 months since start of study treatment, and every 12 weeks (±7 days) thereafter until documented disease progression, initiation of new anticancer therapy or discontinuation of patient from overall study participation (eg, death, patient's request, lost to follow-up), whichever occurs first.

- Overall survival (OS) is defined as the time from the date of first dose of study treatment to the date of death due to any cause.

  Patients will be followed for survival after completion of treatment every 2 months. Contact can be via clinic visit or phone call to patient/caregiver for a minimum of 2 years post-treatment.

Randomized Phase 2 Portion

- EFS assessed by the BIRC

  EFS is defined as the time from date of randomization to the date of first documentation of any of the following events:

  - Objective progressive disease per RECIST v1.1 (including progressive disease without achieving a response or recurrence following a response)
  - Diagnosis of secondary malignancy
  - Death due to any cause

- OR as assessed by investigator using RECIST v.1.1

  OR is defined as a BOR of confirmed CR or PR according to RECIST v1.1 from the date of randomization until the date of the first documentation of PD. Both CR and PR must be confirmed by repeat assessments performed ≥ 4 weeks after the criteria for response are first met. For patients with EWS who have bone marrow disease at study entry, bone marrow aspirates and biopsies are required to confirm CR if on-treatment CR is indicated per RECIST v1.1.

- PET-CT/MRI response after Cycle 4 compared to objective response on MRI/CT

  PET-CT or PET-MRI, if performed at baseline, should be repeated at completion of Cycle 4 (not mandatory). Lesion level FDG avidity (uptake intensity) compared to baseline will be recorded as Resolved, Improved, Stable, Increased, or Not Evaluable. New hot spot (not present at baseline) identified with FDG-PET at completion of Cycle 4 will be recorded as First appearance. Increased FDG avidity alone on PET-CT or PET-MRI without an associated increase in the size or number of lesions must not be used to define disease progression. Objective response on MRI/CT at completion of Cycle 4 will be assessed per RECIST v.1.1.

- Progression-free survival (PFS) based on investigator assessment

  PFS is defined as the time from the date of randomization to the date of the first documentation of objective PD as per RECIST v.1.1 or death due to any cause, whichever occurs first.

- Overall survival (OS)
  OS is defined as the time from the date of randomization to the date of death due to any cause.

3.2.3. Pharmacokinetic Parameters of Palbociclib, IRN, TMZ, TOPO, and CTX: Phase 1 Portion and Tumor-Specific Cohort (if opened)

- TOPO (and active metabolite) PK: MD - $C_{ss,max}$, T_{max}, $AUC_{ss,\tau}$, $C_{ss,trough}$, and CL/F, as data permit.

- CTX (and active metabolite) PK: MD - $C_{ss,max}$, T_{max}, $AUC_{ss,\tau}$, $C_{ss,trough}$, and CL/F, as data permit.

Phase 1 Portion, Tumor-Specific Cohort (if opened) and Randomized Phase 2 Portion

For the Phase 2 portion, PK samples will be collected from patients randomized to palbociclib in combination with IRN and TMZ treatment arm only. No PK samples are required for patients randomized to IRN and TMZ chemotherapy alone arm.

- Palbociclib PK: multiple dose (MD; assuming steady state is achieved) - $C_{ss,max}$, T_{max}, $AUC_{ss,\tau}$, $C_{ss,trough}$, and CL/F, as data permit.

- IRN (and active metabolite, SN-38) PK: MD - $C_{ss,max}$, T_{max}, $AUC_{ss,\tau}$, $C_{ss,trough}$, and CL/F, as data permit.

- TMZ PK: MD - $C_{ss,max}$, T_{max}, $AUC_{ss,\tau}$, $C_{ss,trough}$, and CL/F, as data permit.

3.2.4. Quality of Life (QoL) Endpoints (Phase 2 Portion Only)

- **Patient reported outcomes (PRO)** are measured using the following questionnaires which will be given to the patient and/or their parent/legal guardian/legally acceptable representative (as applicable) at baseline based on patient age at randomization and after completion of Cycles 2 and 4 study treatment and preferably immediately at pre-dose on Day 1 of Cycles 3 and 5:
    - PROMIS Early Childhood Parent Report Scale v1.0 - Global Health 8a instrument is completed by parents for their child ages ≥2 to <5 years old. It assesses a child's overall physical, mental, and social health. It consists of 8 items that produce one global health score with each item having 5 levels of responses (1=Poor, 2=Fair, 3=Good, 4=Very good, and 5=Excellent).The measure is generic, not disease-specific.
    - PROMIS Parent Proxy Profile v2.0 – Profile-25 instrument is intended for parents serving as proxy reporters for their child ages ≥5 to ≤17 years old.
    - PROMIS Pediatric Profile v2.0 – Profile-25 instrument is intended for pediatric self-report for ages ≥8 to ≤17 years old.

      The Profile-25 is a collection of short forms containing 4 items from 6 PROMIS domains (Depressive Symptoms, Anxiety, Physical Function-Mobility, Pain Interference, Fatigue, and Peer Relationships) with each item having 5 levels of responses (1=Never, 2=Almost never, 3=Sometimes, 4=Often, and 5=Almost always) along with a single item on Pain Intensity with 11 levels of responses from 0 (No pain) to 10 (Worst pain you can think of). The Profile-25 is generic rather than disease-specific and assesses all domains over the past 7 days.

- PROMIS–29 Profile v2.1 instrument is intended for adults ages ≥18 years old. It is a collection of short forms containing 4 items from 7 PROMIS domains (Depression, Anxiety, Physical Function, Pain Interference, Fatigue, Sleep Disturbance, and Ability to Participate in Social Roles and Activities) with each item having 5 levels of responses (1=Unable to do, 2=With much difficulty, 3= With some difficulty, 4=With little difficulty, and 5=Without any difficulty) along with a single item Pain Intensity with 11 levels of responses from 0 (No pain) to 10 (Worst pain you can think of). The Profile-29 is generic rather than disease-specific and assesses all domains over the past 7 days except for Physical Function which has no timeframe specified.

- **Days of hospitalization**

  Hospitalization during study includes all planned and unplanned inpatient admissions, regardless the reason (outpatient visits, attendance for transfusions, blood tests and study related procedures etc.), must be captured on the CRF and, if precipitating events meet criteria as per protocol (Sections 8.2.3 and 8.2.4), reported as SAEs. The reporting period for hospital attendance starts with Cycle 1 Day 1 and continues through out study treatment period. The last study visit to report hospitalization is the EOT visit. Any hospitalizations that occurred between the last dose of study treatment and prior to EOT visit would be recorded on the HOS eCRF page at the EOT visit.

3.3. Exploratory Endpoints

- Including but not limited to, abnormalities in P16 and RB genes in pre-treatment archival/de novo biopsy tumor tissue that may be relevant to the potential genomic and biological mechanism of sensitivity and resistance to study treatment.

- Palbociclib exposure and correlation with PD markers (thymidine kinase [TK]-1 activity) in blood.

- Somatic gene mutations and/or germline genetic alterations in patient blood evaluated by DNA sequencing as well as the potential association between the genomic alterations and treatment response.

- Palatability of the palbociclib oral solution formulation as assessed by a questionnaire.

- Bone metabolism and growth, including linear growth, bone age, bone mineral density, physical signs of pubertal maturation (Tanner stage), and hormones associated with growth and pubertal development (insulin-like growth factor 1 [IGF-1], luteinising hormone [LH], follicle-stimulating hormone [FSH], estradiol for girls, testosterone for boys), and a marker of bone formation and bone resorption (bone alkaline phosphatase and C-telopeptide cross-links) (Phase 2 portion only).

3.4. Baseline Variables

3.4.1. Investigational Product, Study Drug, Study Treatment and Baseline Definitions

In this study, the investigational products (IP) are palbociclib, IRN, TMZ, TOPO, and CTX.

Study drug refers to palbociclib, IRN, TMZ, TOPO, or CTX.

Phase 1 Portion

Study treatment (or treatment group) refers to one of the following for each treatment combination.

Combination of palbociclib with IRN and TMZ

Palbociclib oral QD, Day 1-14 + IRN 50 mg/m^2 IV QD, Day 1-5 + TMZ 100 mg/m^2 oral or IV QD, Day 1-5, in a 21-day cycle:

Dose Escalation Part:

- DL-1: Palbociclib 40 mg/m^2
- DL1 (starting dose): Palbociclib 55 mg/m^2
- DL2: Palbociclib 75 mg/m^2
- DL3: Palbociclib 95 mg/m^2
- DL4: Palbociclib 115 mg/m^2

Dose Expansion Cohort:

Based on the dose escalation part of the combination of palbociclib with IRN and TMZ, the MTD for palbociclib has been determined as 75 mg/m^2.

- Dose Expansion: Palbociclib 75 mg/m^2 (MTD)
- Dose Expansion + DL2: Palbociclib 75 mg/m^2, including the 6 patients from Dose Escalation DL2 (ie, patients treated at the MTD of 75 mg/m^2 from the dose escalation part)

Tumor-Specific Cohort(s):

- Palbociclib 75 mg/m^2 (RP2D), for each of specific tumor types (if opened)

Combination of palbociclib with TOPO and CTX

Palbociclib oral QD, Day 1-14 + TOPO 0.75 mg/m^2 IV QD, Day 1-5 + CTX 250 mg/m^2 IV QD, Day 1-5, in a 21-day cycle:

Dose Determination Part:

- DL-1: Palbociclib 55 mg/m^2
- DL1 (starting dose): Palbociclib 75 mg/m^2

Dose Expansion Cohort:

Based on the dose determination part of the combination of palbociclib with TOPO and CTX, the RP2D for palbociclib has been determined as 75 mg/m^2.

- Dose Expansion: Palbociclib 75 mg/m^2 (RP2D)
- Dose Expansion + DL1 (equivalent to Total): Palbociclib 75 mg/m^2, including the 6 patients from Dose Determination DL1 (ie, patients treated at the potential RP2D of 75 mg/m^2 from the dose determination part)

Tumor-Specific Cohort(s)

- Palbociclib RP2D, for each specific tumor type (if opened)

Randomized Phase 2 Portion

The RP2D for palbociclib administered orally QD on Days 1-14 of 21-day cycle was confirmed to be 75 mg/m^2 in combination with IRN at 50 mg/m^2 (IV) daily dose and TMZ at 100 mg/m^2 daily dose on Days 1-5.

Study treatment (or treatment arm) refers to one of the following:

- Palbociclib + IRN +TMZ
- IRN + TMZ

Start and end dates of study treatment

The date of first dose of study treatment in a combination is the earliest date of the first non-zero dose date for the study drugs in the combination.

The date of last dose of study treatment in a combination is the latest date of the last non-zero dose date of the study drugs in the combination.

Definition of baseline for safety analyses

The last available assessment prior to the start of study treatment is defined as “baseline” value or “baseline” assessment for safety analyses. If an assessment is planned to be performed prior to the first dose of study treatment in the protocol and the assessment is performed on the same day as the first dose of study treatment, it will be assumed that it was performed prior to study treatment, if assessment time point is not collected or is missing. If assessment time points are collected, the observed time point will be used to determine pre-dose on study day 1 for baseline calculation. Unscheduled assessments will be used in the determination of baseline. However, if time is missing, an unscheduled assessment on study day 1 will be considered to have been obtained after study treatment administration. No windowing will be applied when defining baseline. Any deviations from the protocol specified window will be documented as protocol deviations.

Patients who start and discontinue treatment on the same day may have two different sets of data collected on study day 1 (one during study and one in the End of Treatment [EOT] visit). Data reported at the EOT visit are not eligible for baseline selection.

If a scheduled pre-dose measurement actually occurred post-dose, then the corresponding measurement will be treated and analyzed similar to an unscheduled post-dose measurement.

ECGs are collected in the study and the baseline for each ECG measurement is the last triplicate mean value assessment performed prior to the first non-zero dose of study drug. Baseline for RR and QT/QTc interval assessments will be derived from the visit where both RR and QT are not missing. QTcB and QTcF will be derived based on RR and QT. The average of the triplicate measurements will be determined after the derivation of the individual parameter at the respective time point.

Definition of baseline for efficacy and PRO analyses

The last measurement prior to first dose of study treatment (for Phase 1 portion) or prior to randomization (for randomized Phase 2 portion) will serve as the baseline assessment for efficacy and PRO analyses unless otherwise specified. If such a value is missing (eg, per protocol the first PRO assessment is planned to occur prior to dosing on Cycle 1 Day 1), the last measurement prior to the first dose of study treatment will be used as the baseline measurement.

Definition of baseline for biomarker analyses

The last measurement prior to first dose of study treatment will serve as the baseline measurement for biomarker analyses. For biomarkers that are planned to be measured on Cycle 1 Day 1, it will be assumed that the measurement was performed prior to study treatment administration, if the assessment time point is not collected or is missing.

3.4.2. Baseline Characteristics

Baseline characteristics (eg, demographics, physical measurements, primary diagnosis, prior anticancer therapies) are described in Section 6.5.1.

For the Phase 2 portion, randomization is stratified by type and time of current disease recurrence (primary refractory or 1st recurrence <2 years vs. 1st recurrence ≥2 years since completion of first line treatment or 2nd or greater recurrence) as recorded in the Interactive Response Technology (IRT) system. Study endpoints (ie, EFS, ORR, PFS, OS) will be stratified by this randomization stratification factor. Other baseline characteristics are not planned to be included as stratification variables or covariates in the statistical models.

3.5. Safety Endpoints

3.5.1. Adverse Events

Safety assessment will consist of monitoring of all AEs, including serious AEs (SAEs), regular monitoring of hematology, clinical chemistry including HbA1c, pregnancy test, and routine monitoring of ECGs (12-lead), and vital signs.

3.5.1.1. Treatment-Emergent Adverse Events (TEAEs)

TEAEs are AEs that occur on or after the first dose of study treatment and up to 35 days after last dose of study treatment.

TEAEs related to study drug as judged by the investigator are considered as treatment-related TEAEs.

3.5.2. Laboratory Data

The hematologic and clinical chemistry laboratory results as characterized by type, frequency, severity as graded by NCI CTCAE version 4.03, and timing.

Laboratory assessment will be assigned to cycles based on the collection date of the sample relative to the start dates of cycles from the study drug administration.

4. ANALYSIS SETS

Data for all patients will be assessed to determine if patients meet the criteria for inclusion in each analysis population prior to releasing the database and classifications will be documented per Pfizer's standard operating procedures.

Only patients who signed informed consent will be included in the analysis sets below.

4.1. Full Analysis Set

Phase 1 Portion and Tumor-Specific Cohort(s)

The full analysis set (FAS) is same as the safety analysis set which includes all enrolled patients who receive at least 1 dose of study drug (ie, palbociclib, TMZ or IRN for palbociclib in combination with TMZ or IRN; palbociclib, TOPO or CTX for palbociclib in combination with TOPO and CTX). The FAS will be the primary analysis set for all safety endpoints, exposure, efficacy endpoints, and patient characteristics.

Randomized Phase 2 Portion

The FAS will include all randomized patients. Patients will be classified according to the study treatment assigned at randomization. The FAS will be the primary population for evaluating all efficacy endpoints, PRO endpoints, and patient characteristics.

4.2. Safety Analysis Set

Phase 1 Portion

The safety analysis set includes all enrolled patients who receive at least one dose of study drug (ie, palbociclib, TMZ or IRN for combination of palbociclib with IRN or TMZ; palbociclib, TOPO or CTX for combination of palbociclib with TOPO and CTX). Patients will be classified according to the study treatment actually received. The full analysis set and the safety analysis set are identical. This will be the primary population for evaluating all safety endpoints, exposure, efficacy endpoints, and patient characteristics.

Randomized Phase 2 Portion

The safety analysis set includes all randomized patients who receive at least one dose of study drug. Patients will be analyzed according to the treatment actually received. The safety analysis set will be the primary population for evaluating safety, treatment compliance, and exposure endpoints.

- If a patient randomized to the palbociclib+IRN+TMZ arm incorrectly received IRN+TMZ for the entire treatment duration, the patient will be classified in the IRN+TMZ arm for the purpose of safety analyses, and vice versa.

- If a patient randomized to the palbociclib+IRN+TMZ arm incorrectly received IRN+TMZ alone not for the entire treatment duration (ie, received at least one dose of palbociclib), the patient will be classified in the palbociclib+IRN+TMZ arm for the purpose of safety analyses.

4.3. Other Analysis Set

4.3.1. Per-Protocol Analysis Set (DLT Evaluable Set)

The following per-protocol analysis sets are applicable to the Phase 1 portion only.

Combination of palbociclib with IRN and TMZ

The per-protocol analysis set (DLT evaluable set) is a subset of the safety analysis set and includes all enrolled patients who receive at least one dose of investigational product (ie, palbociclib, IRN or TMZ) and who do not have major treatment deviations during the first cycle of treatment. Patients with major treatment deviations in the first cycle are not evaluable for DLT or the MTD assessment, unless the treatment deviation is due to a DLT or other treatment-related toxicity, and will be replaced as needed to permit MTD estimation. Major treatment deviations include:

- Administration of less than 80% of the planned dose of palbociclib, IRN, and/or TMZ in Cycle 1 for reasons other than treatment-related toxicity, ie, patients who receive less than 12 doses, 4 doses, and/or 4 doses of prescribed palbociclib, IRN, and/or TMZ, respectively, in Cycle 1 for reasons other than treatment-related toxicity (eg, disease progression, withdrawal of consent) will be considered not evaluable for estimating the MTD and will be replaced;

- Administration of more than 120% of the planned dose of palbociclib, IRN, and/or TMZ in Cycle 1, ie, patients who receive more than 16 doses, 6 doses, and/or 6 doses of prescribed palbociclib, IRN, and/or TMZ, respectively, in Cycle 1 will be considered not evaluable for estimating the MTD and will be replaced;

- Protocol-mandated prophylaxis in Cycle 1 is not administered. Note that patients who did not receive mandated antidiarrheal prophylaxis but did not experience diarrhea that meets DLT criteria will be considered DLT evaluable.

Combination of palbociclib with TOPO and CTX

The per protocol analysis set (DLT evaluable set) includes all enrolled patients who receive at least 1 dose of investigational products (ie, palbociclib, TOPO or CTX) and who do not have major treatment deviations during the first cycle of treatment. Patients with major treatment deviations in the first cycle are not evaluable for DLT or the decision of dose determination, unless the treatment deviation is due to a DLT or other treatment-related toxicity, and will be replaced. Major treatment deviations include:

- Administration of less than 80% of the planned dose of palbociclib, TOPO, and/or CTX in Cycle 1 for reasons other than treatment-related toxicity, ie, patients who receive less than 12 doses, 4 doses, and/or 4 doses of prescribed palbociclib, TOPO, and/or CTX, respectively, in Cycle 1 for reasons other than treatment-related toxicity (eg, disease progression, withdrawal of consent) will be considered not evaluable for the decision of dose determination and will be replaced;

- Administration of more than 120% of the planned dose of palbociclib, TOPO, and/or CTX in Cycle 1, ie, patients who receive more than 16 doses, 6 doses, and/or 6 doses of prescribed palbociclib, TOPO, and/or CTX, respectively, in Cycle 1 will be considered not evaluable for the decision of dose determination and will be replaced;

- Protocol-mandated G-CSF in Cycle 1 is not administered. Note that patients who did not receive mandated G-CSF prophylaxis but did not experience neutropenia that meets DLT criteria will be considered DLT evaluable.

4.4. PK Analysis Set

The PK parameter analysis set is a subset of the safety analysis set who have sufficient information to estimate at least 1 of the PK parameters of interest.

The PK concentration analysis set is a subset of the safety analysis set who have at least 1 reported analyte concentration.

4.5. Biomarker Analysis Set

The biomarker analysis set is a subset of the safety analysis set who have at least 1 baseline biomarker assessment. Analysis sets will be defined separately for blood-based (serum, plasma) and tumor tissue-based (archival tumor tissue, de novo tumor biopsies) biomarkers as appropriate.

5. GENERAL METHODOLOGY AND CONVENTIONS

5.1. Hypotheses and Decision Rules

5.1.1. Hypotheses and Sample Size Determination

Phase 1 Portion

There is no formal hypothesis testing in the Phase 1 portion of the study but rather the focus is on estimation of key summary parameters.

Dose Escalation Part/Dose Determination Part

The exact sample size of the rolling 6 design in the dose escalation part of the study cannot be determined in advance because of the dynamic feature of the design. Each combination of treatment will include a minimum of 6 DLT evaluable patients treated at the MTD or the potential RP2D. The number of patients to be enrolled in the study will depend upon the number of DLTs observed which will in turn determine the number of patients at each dose level and the number of dose levels explored.

The maximum sample size for the dose escalation part for combination of palbociclib with IRN and TMZ is 24 patients (DLT evaluable), assuming that all the 4 dose levels are evaluated and require 6 patients each before the MTD is defined.

The maximum sample size for the dose determination part for combination of palbociclib with TOPO and CTX is 12 patients (DLT evaluable), assuming that 6 patients are evaluated

at the MAD and dose de-escalation proceeds to the next lower dose level with minimum of 6 evaluable patients.

The probability of detecting at least 1 patient with a DLT in 6 evaluable patients is 0.469, 0.738, 0.882, 0.953, and 0.984, when the true underlying DLT rates are 10%, 20%, 30%, 40%, and 50%, respectively. A sample size of 24 patients (at any dose) enables detection of at least 1 AE with a probability of 0.214, 0.708, and 0.920, when the true underlying risk is 1%, 5%, and 10%, respectively.

Dose Expansion Cohorts

A dose expansion cohort with a sample size of 12 patients will allow for estimates of any binary variable (eg, OR rate) with the width of the 95% CI no wider than 0.58 using the exact method and no more than 0.50 using the Wilson method.

A sample size of 12 patients enables detection of at least 1 AE with a probability of 0.114, 0.460, and 0.718, when the true underlying risk is 1%, 5%, and 10%, respectively.

The study will continue to enroll patients until palbociclib PK data (from either treatment combination) are available from a minimum of 6 PK-evaluable patients in each of the following age groups: ≤6 years old, >6 years and <12 years old, and ≥12 years and <18 years old. The expansion cohorts may remain open until the minimum number of PK-evaluable patients have been fulfilled by age group. However, the decision of the RP2D can be based on the minimum of the 12 DLT evaluable patients who complete Cycle 1 DLT evaluation.

Tumor-Specific Cohort(s)

For each treatment combination, once the RP2D is confirmed, other tumor-specific cohorts (maximum of 21 patients each) may be opened if 2 or more patients from the dose escalation part and/or dose expansion cohort for combination of palbociclib with IRN and TMZ or the dose determination part and/or dose expansion cohort for combination of palbociclib with TOPO and CTX show a confirmed OR in the specific tumor type.

Given limited data and heterogeneity of the response rates with either combination across various pediatric tumor types, sample size calculation was driven by the null hypothesis that the ORR is 33% below which further investigation for the combination would not be warranted vs. an alternative hypothesis that the ORR is 58% based on the most likely tumor cohorts to open (eg, neuroblastoma, rhabdomyosarcoma, rhabdoid tumor, medulloblastoma). A modified Simon's 2-stage optimal design (Mander AP, 2010)[8] will be used for each opened tumor-specific cohort with maximum of 21 patients. Under the null hypothesis of a true ORR that does not exceed 33%, this design will control 1-sided type I error to be approximately 0.10. If the true ORR is at least 58%, type II error rate will be no more than 0.20 (ie, at least 80% power).

<u>Randomized Phase 2 Portion</u>

The following statistical hypothesis will be tested to address the primary objective:

H_0: $HR_{EFS} \geq 1$ vs. H_1: $HR_{EFS} < 1$

where HR_{EFS} represents the hazard ratio (palbociclib+IRN+TMZ arm vs. the IRN+TMZ arm) of EFS.

The sample size for the Phase 2 portion of the study is determined based on the assumptions that median EFS for patients receiving IRN and TMZ chemotherapy alone in the treatment of recurrent or refractory EWS is 4.7 months (McCabe MG, 2020)[9] and an improvement to median EFS of 8.1 months (corresponding to a HR=0.58) for palbociclib in combination with IRN and TMZ is considered clinically significant. A total of approximately 50 EFS events are required in the 2 treatment arms based on a 2:1 randomization to have at least 80% power to detect a HR of 0.58 in favor of palbociclib plus IRN and TMZ arm using a 1-sided log-rank test at a significance level of 0.20. One interim futility analysis for EFS is planned at 60% of total events expected (~30 events). A group sequential design with a Rho family (1.0) β-spending function is used to determine the futility boundary. Assuming a 5% and 10% dropout rates by 12 months on the palbociclib combination arm and chemotherapy alone arm, respectively, a non-uniform accrual accomplished over an approximately 25-month period and follow-up for about 6 months after the last patient is randomized, a total sample size of approximately 75 patients (~50 patients in the palbociclib+IRN+TMZ arm and ~25 patients in the IRN+TMZ arm) is required. It is expected approximately 75% of the patients enrolled will be pediatric patients (age ≥2 and <18 years).

5.1.2. Statistical Decision Rules

Phase 1 Portion

Dose Escalation Part

The dose escalation part for combination of palbociclib with IRN and TMZ follows the rolling 6 design allowing for accrual of 2 to 6 patients concurrently onto a dose level. Decisions as to which dose level to enroll a patient will be based on the number of patients currently enrolled and evaluable, the number of patients experiencing a DLT, and the number of patients still at risk of developing a DLT at the time of new patient entry.

Dose escalation occurs if 0 of 3-6 or no more than 1 of 6 evaluable patients experience a DLT while being treated at a dose level. If 2 of 2-6 patients experience DLTs the dose is declared too toxic and the dose level will be de-escalated; otherwise the new patient will enter at the same dose level (Table 4). Accrual to the study will only be suspended when awaiting safety data from 6 patients. When patients are non-evaluable for DLTs, they will be replaced with the next available patient if escalation or de-escalation rules are not fulfilled at the time the next available patient is ready to enter onto the study. Dose re-escalation and intra-patient dose escalation are not permitted on this study.
Table 4 below provides the decision rules for enrolling a patient at the current dose level, at an escalated dose level, at a de-escalated dose level, or whether the cohort is suspended to accrual.

The MTD is empirically defined as the highest dose level at which <33% of a minimum of 6 evaluable patients experience a DLT during Cycle 1 of treatment, when at least 33% of patients experience a DLT at the next highest dose level. If the highest dose level proposed for the study is deemed to be safe, ie, 0 or 1 DLT from 6 patients in the highest dose level, then the MTD will not have been reached and PK analysis will be used to define the dose

selected for the dose expansion cohort or consideration may be given to investigate additional dose levels.

Table 4. Decision Rules in Rolling 6 Design for Combination of Palbociclib with IRN and TMZ

# of Patients	# of Patients with DLT	# of Patients with Data Pending	Decision for next patient
2	2		De-escalate*
	else		Same dose level
3	≥2		De-escalate*
	0	0	Escalate**
	else		Same dose level
4	≥2		De-escalate*
	0	0	Escalate**
	else		Same dose level
5	≥2		De-escalate*
	0	0	Escalate**
	else		Same dose level
6	≥2		De-escalate*
	≤1	0	Escalate**
	0	1	Escalate**
	else		Suspend

Note: In the event that the MTD is not reached, PK analysis will be used to define the dose for dose expansion or additional dose levels to be tested.
*If a minimum of 6 evaluable patients have been treated at next lower dose level, the MTD has been defined;
**If final dose level has been reached with <33% of minimum of 6 evaluable patients experiencing a DLT, the MTD will not have been reached and the highest dose may proceed to dose expansion.

Dose Determination Part

The dose determination part for combination of palbociclib with TOPO and CTX follows a modified rolling 6 design to guide dose de-escalation and to determine the dose proceeding to dose expansion. The design allows for accrual of 2 to 6 patients concurrently onto a dose level. The starting dose will be the MTD determined from the dose escalation part with the IRN and TMZ combination. This dose is considered the MAD and no further dose escalation is planned.

If <33% of a minimum of 6 evaluable patients experience a DLT at the MAD, the MAD will proceed to dose expansion to further confirm the RP2D. If ≥33% of evaluable patients at the MAD experience a DLT, the dose will be de-escalated to the next lower dose level and additional patients will be enrolled at this dose level. If <33% of a minimum of 6 evaluable patients experience a DLT, this dose level will proceed to dose expansion to be confirmed as the RP2D.

Table 5 below provides the decision rules for enrolling a patient at the current dose level and at a de-escalated dose level, or whether the cohort is suspended to accrual.

Table 5. Decision Rules in Rolling 6 Design for Combination of Palbociclib with TOPO and CTX

# of Patients	# of Patients with DLT	# of Patients with Data Pending	Decision for next patient
2	2		De-escalate*
	else		Same dose level
3	≥2		De-escalate*
	else		Same dose level
4	≥2		De-escalate*
	else		Same dose level
5	≥2		De-escalate*
	else		Same dose level
6	≥2		De-escalate*
	≤1	0	Proceed**
	0	1	Proceed**
	else		Suspend

*If a minimum of 6 evaluable patients have been treated at the next lower dose level, the dose will proceed to dose expansion.
**If <33% of minimum of 6 evaluable patients treated at the MAD experiencing a DLT, the MAD will proceed to dose expansion.

Dose Expansion Cohorts

The RP2D is the dose chosen for further investigation based on the results of dose escalation part/dose determination part and dose expansion cohorts of the study. If the MTD/potential RP2D proves to be clinically feasible for long-term administration in a reasonable number of patients, then this dose usually becomes the RP2D. Further experience with the MTD/potential RP2D based on other safety, clinical activity, PK, and PD data may result in a RP2D dose lower than the MTD/potential RP2D.

Tumor-Specific Cohort(s)

A modified Simon's optimal 2-stage design (Table 6) will be used for each of the opened tumor-specific cohorts. In the first stage, 7 patients will be enrolled in each tumor-specific cohort. Patients with confirmed ORs (responders) along with any other patients (non-responders) with the same tumor type from dose escalation part and/or dose expansion cohort for combination of palbociclib with IRN and TMZ, or from dose determination part and/or dose expansion cohort for combination of palbociclib with TOPO and CTX will count towards the 7 patients in the first stage. If there are only 2 confirmed ORs in these 7 patients, the cohort will be stopped for ineffectiveness, (ie, none of the additional enrolled patients had a confirmed OR). If there are 6 or more confirmed ORs in these 7 patients, the cohort will be stopped for effectiveness. Otherwise, 14 additional patients will be enrolled in the second stage for a total of 21 patients in each tumor-specific cohort. The null hypothesis will be rejected if 10 or more confirmed ORs are observed in the total of 21 patients.

Table 6. Modified Simon's Optimal 2-Stage Design Decision Rules for Tumor-Specific Cohorts

	Cumulative Number of Responses[1]	Decision
Stage 1: Enroll a total of 7 patients[2]	2	Terminate enrollment – agent ineffective
	3-5	Inconclusive – proceed to Stage 2
	6[3]	Terminate enrollment – agent effective
Stage 2 Enroll 14 additional patients	9 or less	Agent ineffective
	10 or more	Agent effective

1. Assumption based on a true ORR of 33% and 58% for IRN/TMZ or TOPO/CTX and palbociclib/IRN/TMZ or palbociclib/TOPO/CTX, respectively.
2. Evaluated patients will include the 2 ORs observed and any other patients with the same tumor type from the dose escalation part/dose determination part and/or dose expansion cohorts of the study.
3. Lower bound of the Wilson's 2-sided 80% confidence interval is greater than 62%.

Randomized Phase 2 Portion

The randomized Phase 2 portion of the study will have 1 formal interim analysis for futility to allow for early stopping of the study due to futility/no signal of activity based on the primary endpoint of investigator-assessed EFS. No interim analysis for efficacy/superiority is planned. The interim futility and final analyses of EFS will be performed in the FAS after the target number of events have occurred as described below:

- An interim futility analysis of EFS at the time when approximately 30 investigator-assessed EFS events (60% of total events expected, ie, 60% of information for EFS) have been documented. A group sequential design with a Rho family (1.0) β-spending function is used to determine the futility boundary (binding). If the 1-sided p-value is >0.4287 (Z >-0.1796) at the time of the interim analysis, the Phase 2 portion of the study will be terminated early for lack of efficacy.

- The final analysis of EFS at the time when approximately 50 investigator-assessed EFS events (100% of information for EFS) have been documented. The nominal 1-sided significance level for the final analysis of EFS is 0.20. There is no penalty for α spending with the interim futility analysis. The study would be considered to meet its primary endpoint if the 1-sided log-rank test (stratified) for EFS based on the FAS is significant at the significance level of 0.20 at the final analysis (with ~50 events).

5.2. General Methods

For the Phase 1 portion of the study, statistical inference will be mainly based on estimation. An overview of the summaries and the tabulations for this study is described in Table 7. Summaries will be presented by the two treatment combinations separately for palbociclib+IRN+TMZ and palbociclib+TOPO+CTX.

Table 7. Study Summaries and Tabulations for Each Treatment Combination - Phase 1 Portion and Tumor-Specific Cohort(s)

Summaries	Analysis set	Dose Escalation Part / Dose Determination Part	Dose Expansion Cohorts	Tumor-Specific Cohort(s)
Baseline characteristics	FAS (safety analysis set)	- by study treatment group[1] - pooled: all patients[2]		- by tumor type - pooled: all patients
Dispositions	FAS (Safety analysis set)	- by study treatment group[1] - pooled: all patients[2]		- by tumor type - pooled: all patients
DLTs (Cycle 1)	Per-protocol analysis set (DLT evaluable set)	- by study treatment group[1] - pooled: all patients[2]		Not applicable
Other safety data, exposure, concomitant medications, non-drug treatment	Safety analysis set	- by study treatment group[1] - pooled: all patients[2]		- by tumor type - pooled: all patients
Efficacy	FAS (safety analysis set)	- by study treatment group[1] - pooled: all patients[2]		- by tumor type - pooled: all patients
PK data	PK analysis set (PK parameter analysis set and PK concentration analysis set)	- by study treatment group[1] (if data permit) - pooled: all patients[3]		- by tumor type - pooled: all patients[3]
Biomarker data (as applicable)	Biomarker analysis set	- by study treatment group[1] (if data permit) - pooled: all patients[2]		- by tumor type - pooled: all patients

[1] Study treatment group specified in Section 3.4 for dose escalation part, dose determination part, dose expansion cohorts and tumor-specific cohorts, respectively.
[2] Including 1) all patients (dose escalation part/dose determination part and dose expansion cohort) treated at the MTD/potential RP2D, and 2) all patients (dose escalation part/dose determination part + dose expansion cohort).
[3] The pooled PK analysis consists of all three portions of the study (ie, including tumor specific cohorts as well) and be presented by three different stratifications: 1) by analyte dose level (all patients), 2) by analyte dose level (all pediatric patients, ie, <18 years old), and 3) by age group (eg, ≤6, >6 - <12, ≥12 and <18 years old).

For the randomized Phase 2 portion of the study, treatment arm refers to one of the following as described in Section 3.4.1 with palbociclib administered at RP2D of 75 mg/m^2.

- Palbociclib + IRN + TMZ
- IRN + TMZ

Endpoints will be summarized based on the analysis sets described in Table 8 by treatment arm, unless otherwise specified.

5.2.1. Analyses for Categorical Data

Categorical variables, including binary variables, will be summarized by frequency counts and percentages. Unless otherwise specified, the calculation of percentages will include the missing category, ie, counts of missing observations will be included in the denominator and presented as a separate category.

5.2.2. Analyses for Continuous Data

Continuous variables will be summarized by descriptive statistics, including the non-missing number, mean, standard deviation (SD), median, and range (minimum, maximum). Select PK parameters will include geometric mean and its associated coefficient of variation (CV%) as well.

Table 8. Statistical Analyses by Analysis Set - Phase 2 Portion

Endpoints	Full Analysis Set	Safety Analysis Set	PK Analysis Set	Biomarker Analysis Set
Baseline Characteristics	✓			
Prior Anticancer Therapies	✓			
Concomitant Medications		✓		
Follow-Up Anticancer Therapies	✓			
Efficacy	✓*			
PRO	✓			
Safety		✓		
Compliance and Exposure		✓		
PK			✓	
Biomarker (as applicable)				✓

* Some secondary efficacy analyses are based on the subset of FAS.

5.2.3. Analyses for Time to Event Data

Time-to-event endpoints (eg, EFS, PFS, OS, DOR) will be summarized using the Kaplan-Meier method and estimated survival curves will be displayed graphically when appropriate. The number of patients at risk over time will be displayed. Median event times and quartiles (25% and 75%) will be estimated along with corresponding 2-sided 95% confidence intervals (CI) which will be calculated according to Brookmeyer and Crowley method. The CIs for the survival function estimates at the specific time points will be derived using the log-log transformation (conftype=loglog default option in SAS Proc LIFETEST) with back transformation to a CI on the untransformed scale. The estimate of the standard error will be computed using Greenwood's formula.

For time-to-event endpoints in the randomized Phase 2 portion of the study, the log-rank test will be used for comparing treatments. Hazard ratios and the associated 2-sided 95% CIs (unless otherwise specified) are estimated by Cox proportional hazards models, including treatment as a covariate. Ties will be handled using the Exact option in SAS (Ties=Exact option in SAS PROC PHREG).

5.2.4. Definition of Study Day

Start day of study treatment is the day of the first dose of study treatment.

The study day for assessments occurring on or after the first dose of study treatment (eg, AE onset, tumor measurement) will be calculated as:

Study day = Date of the assessment/event - start date of study treatment + 1.

The study day for assessments occurring prior to the first dose of study treatment (eg, baseline characteristics, medical history) will be negative and calculated as:

Study day = Date of the assessment/event - start date of study treatment.

The study day will be displayed in all relevant data listings.

5.2.5. Definition of On-Treatment Period

Safety endpoints will be summarized based on the on-treatment period unless otherwise specified. On-treatment period is defined as the time from the first dose of study treatment to up to 35 days after last dose of study treatment.

Safety data collected outside the on-treatment period as described above will be listed and flagged in listings but not summarized.

5.2.6. Definition of Start of New Anticancer Therapy

Start date of new anticancer therapy (drug, radiation, surgery) is used for censoring in the sensitivity analysis for the primary endpoint of EFS (Section 6.2.2).

The start date of new anticancer therapy is the earliest date after randomization among the following:

- Start date of anticancer drug therapy recorded in the On Study and/or Follow-up Cancer Therapy eCRF pages.
- Start date of radiation therapy recorded in On Study and/or Follow-up Radiation Treatment' eCRF pages.
- Date of Surgery recorded in On Study and/or Follow-up Cancer Surgery eCRF pages when 'Outcome of Procedure' = 'Resected' or 'Partially Resected'.

5.2.7. Standard Derivation and Reporting Conventions

The following conversion factors will be used to convert days into weeks, months or years:

1 week = 7 days, 1 month = 30.4375 days, 1 year = 365.25 days.

Demographics and physical measurements:

- Age at screening
    - The first day of the month will be used to impute the missing date given only month and year of birth are collected:
      Age (years) = (date of given informed consent – date of birth + 1) / 365.25

The integer part of the calculated age will be used for reporting purposes.

- BMI (kg/m^2) = weight (kg)/[height (m)]^2

For reporting conventions, mean and median should generally be displayed one more decimal place than the raw data and standard deviation should be displayed to two more decimal places than the raw data. Percentages will be reported to one decimal place.

If the day of the month is missing for a start date used in a calculation, the first day of the month will be used to impute the missing date.

5.2.8. Unscheduled Visits

Generally, data collected at unscheduled visits will be included and analyzed for both safety and efficacy analyses (eg, maximum CTCAE grade, defining date of progression/censoring, BOR) in the same fashion as the data collected at scheduled visits except where otherwise specified. Descriptive statistics (mean, SD, median, minimum, maximum) by nominal visit or time point for safety endpoints such as laboratory measurements and vital signs will include only data from scheduled visits.

5.2.9. Adequate Baseline Tumor Assessment

Adequate baseline is defined using the following criteria:

- All baseline assessments must be within 21 days (+3 days) prior to and including Cycle 1 Day 1 date (Phase 1 portion and tumor-specific cohorts).
- All baseline assessments must be within 21 days (+3 days) prior to and including the date of randomization or within 3 days post-randomization prior to and including the date of first dose of study treatment (Phase 2 portion).
- All documented lesions must have non-missing assessments (ie, non-missing measurements for target lesions and non-missing lesions assessment status at baseline for non-target lesions).

5.2.10. Adequate Post-Baseline Tumor Assessment

An adequate post-baseline assessment is defined as an assessment after start of study treatment where a response of CR, PR, MR, SD, or PD can be determined (see Section 6.1.2). Time points where the response is not evaluable (NE) or no assessment was performed will not be used for determining the censoring date in time-to-event analysis.

5.3. Methods to Manage Missing Data

Unless otherwise specified, all data will be evaluated as observed, and no imputation method for missing values will be used. In all patient data listings imputed values will be presented and flagged.

Missing statistics, when they cannot be calculated, should be presented as ‘NA’ (not applicable), eg, if N=1, the measure of variability (SD) cannot be computed and should be presented as ‘NA’.

5.3.1. Missing or Partial Dates

In general, imputation methods for partial dates as follows unless otherwise specified:

- If the day of the month is missing for a start date used in a calculation, the first day of the month will be used to impute the missing date.

- If both the day and month are missing for a start date, the first day of the year, January 1st, is used.
- For stop dates, the last day of the month, or last day of the year is used if the day or day and month are missing, respectively.
- If the date is completely missing, no imputation will be performed.

These rules are used unless the calculations result in negative time durations (eg, date of onset cannot be prior to Day 1 date). In these cases, the dates resulting in 1-day duration will be used.

5.3.1.1. Adverse Events

- If the AE start date is missing completely, then the start date will be imputed by the start of study treatment.

- If only the day of the AE start date is missing, but the month and year are equal to the start of study treatment, then the AE start date will be imputed by the start of study treatment.

- If both the day and month of the AE start date are missing but the year is equal to the start of study treatment, then the start date will be imputed by the start of study treatment.

- In all other cases the missing start day or missing start month will be replaced by 1.

- Incomplete stop date will be imputed by the last day of the month (if day is missing only), if not resulting in a date later than the date of death. In the latter case the date of death will be used to impute the incomplete stop date.

- In all other cases the incomplete stop date will not be imputed. If stop date of AE is after the date of cutoff, outcome of AE is ongoing as of the cutoff.

5.3.1.2. Exposure

No imputation will be done for first dose date. Date of last dose of study drug, if unknown or partially unknown, will be imputed as follows:

- If the date of last dose of study drug is completely missing and there is no Disposition Treatment eCRF page and no death date, the patient should be considered to be ongoing and use the cutoff date for the analysis as the last dose date;

- If the date of last dose of study drug is completely or partially missing and there is either a Disposition Treatment eCRF page or a death date available (within the cutoff date), then the imputed last dose date is:

  = 31DECYYYY, if only Year is available and Year < Year of min (EOT date, death date)

= Last day of the month, if both Year and Month are available and Year = Year of min (EOT date, death date) and Month < the month of min (EOT date, death date) = min (EOT date, death date), for all other cases.

5.3.1.3. Date of Last Contact

The date of last contact will be derived for patients not known to have died at the analysis cutoff using the latest complete date (non-imputed) among but not limited to the following:

- All patient assessment dates (eg, laboratory assessments, PK, vital signs, performance status, ECG, tumor assessments, concomitant medication, and non-drug treatments)
- Start and end dates of anticancer therapies
- AE start and end dates
- Study drugs start and end dates
- Randomization date (Phase 2 portion)
- Completion dates for PRO questionnaires
- Withdrawal of consent date
- Date of last known to be alive collected on the "Survival Follow-Up" eCRF page where "Subject Remains in Follow-up" or "Subject No Longer Being Followed for Survival" is selected
- Date of completion/discontinuation on the Disposition eCRF pages including withdrawal of consent date (do not use if reason for discontinuation is lost to follow-up).

Only dates associated with actual examinations of the patient will be used in the derivation. Dates associated with a technical operation unrelated to patient status such as the date a blood sample was processed will not be used. Assessment dates after the cutoff date will not be applied to derive the last contact date.

5.3.1.4. Death Date

Missing or partial death dates will be imputed based on the last contact date:

- If the date is missing it will be imputed as the day after the date of last contact.

- If the day or both day and month is missing, death will be imputed to the maximum of the full (non-imputed) day after the date of last contact and the following:

    - Missing day: 1st day of the month and year of death
    - Missing day and month: January 1st of the year of death

5.3.1.5. Tumor Assessments

If baseline tumor assessment is inadequate the patient cannot be assessed for response.

If there are multiple scan dates associated with an evaluation, ie, radiological assessments occur over a series of days rather than the same day, the choice of date of assessment could impact the date of progression and/or date of response. Therefore, the earliest of the scan dates associated with the evaluation will be used as the date of assessment.

If all measurement dates for an evaluation have no day recorded, the 1st of the month is used.

If the month is not completed, for any of the investigations for an evaluation, the respective assessment will be considered to be at the date which is exactly between the previous and the following assessment. If both a previous and following assessments are not available, this assessment will not be used for any calculations.

If measurements for one or more target lesions are missing for an evaluation and disease does not qualify as progression, the objective status for that evaluation is NE.

If non-target disease was not assessed, then objective status cannot be a CR even if all target disease has disappeared. Otherwise, missing non-target disease assessments do not necessarily affect response determination. Such cases will be reviewed carefully.

If a lesion measurement is missing because it is documented as too small to measure, the value 5 mm will be assigned, and objective status calculated accordingly.

A target/non-target lesion that is irradiated or excised due to concurrent surgery for local control post-baseline is no longer measurable and should be assessed as Not Evaluable. Increase or reappearance of such lesions will constitute progression.

In the assessment of OR, patients who do not have on study repeat radiographic assessments will be counted as non-responders.

5.3.1.6. Missing Data in Time-to-Event Endpoints

For all time to event endpoints unless otherwise specified, if the day of the month is missing for any date used in a calculation, the first day of the month will be assigned. However, if the month and the year of the event are same as the month and year when the patient was randomized, then the date of randomization will be assigned. The time of a time-to-event endpoint cannot be negative.

Dropouts without treatment are defined as patients who have been randomized in the study but withdraw their consent or are withdrawn by the investigator from the study immediately thereafter but prior to first administration of study drug. These patients will be included in the FAS with censorship defined in Section 6.

In case of missing disease assessments before the last disease assessment or before death, EFS and PFS events will be defined based on the actual disease assessment and not the scheduled disease assessments.

All other missing values will be reported explicitly.

5.3.2. Missing QTc Data

For QTc analysis, no values will be imputed for missing data except for averaging of triplicate measurements. If one or two of the triplicate measurements for an ECG parameter are missed, the average of the remaining two measurements or the single measurement can be used in the analyses. If all triplicate measurements are missing at a time point for an ECG parameter, no values will be imputed for this time point and no analyses related to this time point will be performed.

5.3.3. Pharmacokinetic Concentrations

Concentrations below the limit of quantification

For all calculations, figures and estimation of individual PK parameters, all concentrations assayed as below the level of quantification (BLQ) will be set to zero. In log-linear plots these values would not be represented. The BLQ values will be excluded from calculations of geometric means and their CIs. A statement similar to "Summary statistics have been calculated by setting concentration values below the lower limit of quantification to zero" should be included as a footnote to the appropriate tables and figures.

Deviations, missing concentrations and anomalous values

In summary tables and plots of median profiles, statistics will be calculated. Concentrations will be set to missing if one of the following cases is true:

- A concentration has been reported as ND (ie, not done) or NS (ie, no sample);
- A deviation in sampling time is of sufficient concern or a concentration has been flagged as anomalous by the clinical pharmacologist or oNCA pharmacometrician.

Summary statistics may not be presented at a particular time point if more than 50% of the data are missing. For analysis of PK concentrations, no values will be imputed for missing data.

5.3.4. Pharmacokinetic Parameters

Actual PK sampling times (and where possible the actual dosing information) will be used in the derivation of PK parameters. If a PK parameter cannot be derived from a patient's concentration data, the parameter will be coded as NC (ie, not calculated). NC values will not be generated beyond the day that a patient discontinues. For calculation of AUCτ, if the 24-hr sample for Cycle 1 Day 5 is missing, imputation of 24-hr sample concentration using reported pre-dose concentrations on subsequent days or cycles may be done, if data permit, as applicable for each analyte. This will only occur when all other planned samples for the serial collection on Cycle 1 Day 5 have been completed and have reportable concentrations.

In summary tables, statistics will be calculated by setting NC values to missing; and statistics may not be presented for a particular treatment if more than 50% of the data are NC. For statistical analyses (ie, analysis of variance), PK parameters coded as NC will also be set to missing.

If an individual patient has a known biased estimate of a PK parameter (due for example to an unexpected event such as vomiting before all the drug is absorbed in the body), or missed doses prior to a collection of PK sample meant to be at steady-state), this will be footnoted in summary tables and will not be included in the calculation of summary statistics or statistical analyses.

5.3.5. Missing PRO Data

There will be no imputation of missing PRO data for analysis purposes.

The age appropriate PROMIS questionnaires will be scored according to their respective scoring manuals. For the 8-item Early Childhood Parent Report Global Health 8a form, all questions must be answered in order to produce a valid score. If a patient has skipped a question, the total score is deemed missing. For other questionnaires, if the answer to a question of any domain is missing, then the total score for that domain is deemed missing.

5.3.6. Missing/Duplicate Biomarker Data

No values will be imputed for missing data. If there is duplicate biomarker data for an assessment that is not expected (eg, the same sample was sent twice to the assay laboratory, or the laboratory reported two assessments for the same sample), the following will occur:

- Numerical biomarker values will be averaged for reporting purposes and the average will be included in the listing of biomarker data along with the original data.
- Non-numerical data (eg, mutations present) will be reviewed by the clinical to determine which data will be included in the analyses, if any; that record will have a flag added to show it was used for analysis purposes.

6. ANALYSES AND SUMMARIES

All the analyses will be presented separately for the Phase 1 portion and randomized Phase 2 portion of the study.

The following analyses will be summarized separately by each treatment combination as shown in Table 7 for the Phase 1 portion and Table 8 for the randomized Phase 2 portion, unless otherwise specified.

6.1. Primary Endpoints

6.1.1. First Cycle DLT (Phase 1 Portion)

The primary endpoint is the occurrence of DLTs in the first cycle (Cycle 1 Day 1 until pre-dose Cycle 2 Day 1) which is used to determine the MTD in the dose escalation part for the combination of palbociclib with IRN and TMZ and to determine the potential RP2D in the dose determination part for the combination of palbociclib with TOPO and CTX.

6.1.1.1. Primary analysis

Analyses of DLTs are based on the per-protocol analysis set (DLT evaluable set). The occurrence of first cycle DLTs will be summarized and listed by dose level and treatment groups. First cycle DLTs by MedDRA PT (including clusters of PTs) will also be summarized by dose level and treatment groups.

6.1.2. Adverse Events (Phase 1 Portion and Tumor-Specific Cohorts)

6.1.2.1. Primary analysis

Safety evaluations for the dose expansion cohorts will be based on the safety analysis set. AEs will be classified using latest version of the MedDRA classification system. Severity of AEs will be graded according to the NCI CTCAE v.4.03. ECG parameters and vital signs during on-treatment period will be summarized.

6.1.3. Objective Response as Assessed by the Investigator for the Dose Expansion Cohort (Phase 1 Portion) and Tumor-Specific Cohort(s)

6.1.3.1. Primary analysis

Analyses of OR will be based on the FAS (safety analysis set). OR will be assessed by the investigator per RECIST v1.1, modified RANO for CNS malignancies or INRC for neuroblastoma.

BOR will be assessed based on reported overall lesion responses at different evaluation time points from the start of study treatment until the first documentation of PD or the start of new anticancer therapy, whichever occurs first. Only tumor assessments performed on or before the start date of any further anticancer therapies will be considered in the assessment of BOR. Clinical deterioration alone will not be considered as documentation of objective PD. Confirmation of response is required ≥4 weeks after initial response is observed.

- **CR** (confirmed): At least two objective statuses of CR ≥4 weeks apart and before first documentation of PD.
- **PR** (confirmed): At least two objective statuses of PR or better (PR followed by PR or PR followed by CR) performed ≥4 weeks apart before first documentation of PD (and not qualifying for a CR). Two PRs separated by one or more SD is considered a confirmed PR as long as the two PR responses are observed ≥4 weeks apart.
- **MR** (INRC only; confirmed): At least two objective statuses of MR or better ≥4 weeks apart and before first documentation of PD.
- **SD**: At least one objective status of SD or better documented ≥2 cycles after treatment start date and before first documentation of PD (and not qualifying for CR or PR).
- **PD**: Documentation of objective status of PD prior to the second planned assessment (ie, ≤4 cycles) after treatment start date (and not qualifying for CR, PR, MR, or SD).
- **Not Evaluable (NE)**: All other cases.

An objective status of PR or SD cannot follow one of CR. A sequence of PR-SD-SD-PD would be a best response of SD if the window for SD definition has been met.

OR is defined as a BOR of confirmed CR or PR according to RECIST v1.1 or modified RANO for CNS malignancies. For neuroblastoma based on INRC, OR is defined as a BOR of confirmed CR, PR, or MR.

ORR will be calculated as the proportion of patients with OR along with the 2-sided 95% CI using the Wilson method (Wilson, 1927[16]; Agresti A,1998[1]) by dose level and across all dose levels. For each tumor-specific cohort, if opened, ORR will be calculated with the 2-sided 80% CI using the Wilson method. A 2-sided Wilson's 95% CI may also be reported.

Patients will be considered non-responders in the assessment of OR if they

- do not have a confirmed CR or PR while on study (confirmation of response is required ≥4 weeks after the criteria for response are first met),
- do not have a post-baseline radiographic tumor assessment,
- receive anticancer therapies other than the study treatments prior to reaching a CR or PR or
- die, or have documented PD prior to reaching a CR or PR.

In addition, the number and percentage of patients with a BOR of confirmed CR or PR, SD, PD, and NE will be tabulated. Reasons for NE will be summarized with the following reasons:

- No adequate baseline assessment
- No post-baseline assessments due to early death (<2 cycles after treatment start date)
- No post-baseline assessments due to other reasons
- All post-baseline assessments have overall response NE
- New anticancer therapy started before first post-baseline tumor assessment
- SD too early (<2 cycles after treatment start date without further evaluable tumor assessments)
- PD too late (after two inadequate assessments or after missing two consecutive planned assessments, ie, >4 cycles after treatment start date)

Special and rare cases where BOR is NE due to both SD of insufficient duration and late PD will be classified as "SD too early".

6.1.4. EFS Based on Investigator Assessment (Phase 2 Portion)

6.1.4.1. Primary analysis

The primary analyses of EFS will be based on the FAS using the strata assigned at randomization.

EFS is defined as the time from date of randomization to the date of first documentation of any of the following events based on investigator assessment:

- Objective progressive disease per RECIST v.1.1 (including progressive disease without achieving a response or recurrence following a response)
- Diagnosis of secondary malignancy
- Death due to any cause

EFS (months) = (date of event or censoring – date of randomization + 1) / 30.4375

Patients last known to be alive and event-free will be censored at the date of the last objective disease assessment that verified lack of an event. In addition,

- Patients with no adequate baseline tumor assessment (including patients with an inadequate baseline assessment) will be censored on the date of randomization with duration of 1 day.

- Patients with no adequate post-baseline tumor assessments will be censored on the date of randomization with duration of 1 day, unless death occurred on or before the second planned tumor assessment (ie, ≤12 weeks after the date of randomization) in which case the death is an event with date on death date.

- Patients with an event after 2 or more missing or inadequate post-baseline disease assessments (including not evaluable assessments due to local control) will be censored at the time of last objective disease assessment documenting no event.

- Patients who are removed from the study (withdrawal of consent, lost to follow-up, etc.) prior to an event will be censored at the date of the last objective disease assessment documenting no event.

A stratified log-rank test (1-sided) will be used to compare EFS between the two treatment arms. The HR for EFS and associated 2-sided 95% CI will be estimated using the stratified Cox's Proportional Hazard model. A 2-sided 60% CI of the HR will also be reported.

Kaplan-Meier estimates will be presented by treatment arm with a summary of associated statistics including the median EFS time and corresponding 2-sided 95% CIs. EFS (%) at specific time points (eg, 3, 6, 12, 18, 24, 30, and 36 months) will be estimated with corresponding 2-sided 95% CIs for each treatment arm using the Kaplan-Meier method.

Number and percentage of patients with each event type and censoring reasons will be presented by treatment arm. Reasons for censoring will be summarized following the category hierarchy shown in Table 9:

Table 9. EFS Censoring Reasons and Hierarchy

Hierarchy	Condition	Censoring Reason
1	No adequate baseline tumor assessment	No adequate baseline assessment
2	Event after 2 or more missing or inadequate post-baseline tumor assessments including not evaluable assessments due to local control	Event after ≥2 missing or inadequate post-baseline assessments
3	No event and withdrawal of consent date ≥ date of randomization	Withdrawal of consent

Table 9. EFS Censoring Reasons and Hierarchy

Hierarchy	Condition	Censoring Reason
4	No event and lost to follow-up in any disposition page	Lost to follow-up
5	No event and no adequate post-baseline tumor assessment	No adequate post-baseline assessment
6	No event and none of the conditions in the prior hierarchy are met	In follow-up without an event

The EFS event or censoring time and the reasons for censoring will also be presented in a patient listing.

Time of Follow-Up for EFS

A Kaplan-Meier plot for EFS follow-up duration will be generated to assess the follow-up time in the treatment arms reversing the EFS censoring and event indicators. Median duration of follow-up will be provided with the 95% CIs.

6.1.4.2. Sensitivity analyses

The following sensitivity analyses will be performed to explore the robustness of the primary analysis results of EFS based on investigator assessment. These analyses are regarded as purely exploratory. The sensitivity analyses will repeat the primary analysis (p-value, HR and 95% CIs) described in Section 6.1.4.1 with the modifications below:

- Unstratified analysis (Sensitivity Analysis 1) – A sensitivity analysis will repeat the primary analysis of EFS described in Section 6.1.4.1 using an unstratified analysis.

- Censoring for initiation of new anticancer therapy prior to an event (Sensitivity Analysis 2) – A sensitivity analysis will be performed in which patients who initiate a new anticancer therapy prior to an EFS event will be censored at the date of last tumor assessment prior to the initiation of the new therapy. Patients who do not have an adequate post-baseline tumor assessment will be censored on the date of randomization unless the patient dies ≤12 weeks after the date of randomization and does not initiate new anticancer therapy, in which case the death is an event. Table 9 will be modified to include "Initiation of new anticancer therapy" as a censoring reason after "No adequate baseline assessment".

- Stratification factor based on CRF data (Sensitivity Analysis 3) – A sensitivity analysis will be performed using the stratification factor based on the data entered by the investigator in the "Primary Diagnosis" eCRF at screening.

- Per-protocol analysis of EFS in patients with molecularly confirmed EWS at diagnosis or at relapse or by central testing (Sensitivity Analysis 4) – A sensitivity analysis will be performed excluding patients who are not confirmed by pathology report or central testing for EWS diagnosis.

- Multivariable Cox regression analysis (Sensitivity Analysis 5) – A sensitivity analysis will be performed using the Cox's Proportional Hazard model adjusting the treatment effect for the stratification factor (based on eCRF) and relevant baseline factors of potential prognostic impact.

- Potential impact of COVID-19 related deaths (Sensitivity Analysis 6) – If >10% of the FAS patients have died due to COVID-19 without a prior event, a cumulative incidence analysis will be performed for EFS (event of interest excluding COVID-19 related deaths) by treating COVID-19 related death as a competing risk event.

Methods for evaluating the validity of model assumptions

The proportional hazards (PH) assumption will be tested using Schoenfeld's residual test (Schoenfeld, 1980[11]; Therneau & Grambsch, 2000[13]). Departure from PH could be evidenced by a p-value <0.05. If there is evidence of large departures from proportional hazards, EFS will also be analyzed based on restricted mean survival time (RMST) differences (Zhang, 2013[17]).

The RMST up to time τ can then be interpreted as the expected survival time restricted to the common follow-up time τ among all patients. To avoid arbitrary selection of the common cut-off τ for both treatment arms, three sets of analyses will be performed:

- τ_1 = minimum of (largest observed survival time for experimental arm, largest observed survival time for control arm).
- τ_2 = minimum of (largest survival event time for experimental arm, largest survival event time for control arm).
- $\tau 3$ = midpoint between $\tau 1$ and $\tau 2$.

The treatment effect between the experimental arm and the control arm will be assessed based on the difference in RMST. The associated 95% CI for the difference in means and 1-sided p-value will be generated.

6.2. Secondary Endpoints

6.2.1. Safety Endpoints

Refer to Section 6.6.

6.2.2. Efficacy Endpoints

The following analyses will be based on the FAS defined in Section 4.

Tumor response will be assessed using RECIST v1.1, modified RANO for CNS malignancies (Phase 1 portion) or INRC for neuroblastoma (Phase 1 portion). Disease assessment related endpoints (OR, DOR, PFS) will be based on investigator assessment. In

the Phase 2 portion, EFS by BIRC assessment will be analyzed as a secondary endpoint. OR, DOR and PFS will be summarized separately based on investigator assessment and BIRC results.

6.2.2.1. EFS assessed by the BIRC

EFS by BIRC assessment will be summarized using the same methodology as described in Section 6.1.4.1 for the primary analysis of EFS by investigator assessment.

A descriptive summary of agreement between investigator and BIRC assessments will be provided based on timing and occurrence of event. Number and percentage of patients with possible outcomes of concordance/discordance for investigator vs. BIRC assessments will be provided for each treatment arm and overall:

- Agreement by investigator and BIRC on timing and occurrence of event
- Agreement on occurrence of event but investigator declares event later than BIRC
- Agreement on occurrence of event but investigator declares event earlier than BIRC
- Event by investigator and no event by BIRC
- No event by investigator and event by BIRC
- Agreement on no event by investigator and BIRC

6.2.2.2. OR

For the Phase 1 portion, OR will be summarized using the methodology described in Section 6.1.3.

For the Phase 2 portion, OR is defined as a BOR of confirmed CR or PR according to RECIST v1.1. The number and percentage of patients with a BOR of confirmed CR, confirmed PR, SD, non-CR/non-PD (applicable only to patients with non-measurable disease at baseline), PD, and NE will be tabulated. OR will be summarized separately based on investigator assessment and BIRC results.

For patients with bone marrow disease at study entry, bone marrow aspirates and biopsies are required to confirm CR if on-treatment CR is indicated per RECIST v1.1 at the confirmatory scan. If there is bone marrow involvement (ie, tumor infiltration >0) at this evaluation, the objective response status will be classified as PR (ie, CR not confirmed).

BOR will be assessed based on reported overall lesion responses at different evaluation time points from the date of randomization until the first documentation of PD or the start of new anticancer therapy, whichever occurs first. Only tumor assessments performed on or before the start date of any further anticancer therapies will be considered in the assessment of BOR. Clinical deterioration alone will not be considered as documentation of objective PD. Confirmation of response is required ≥4 weeks after initial response is observed.

- **CR** (confirmed): At least two objective statuses of CR ≥4 weeks apart and before first documentation of PD.

- **PR** (confirmed): At least two objective statuses of PR or better (PR followed by PR or PR followed by CR) performed ≥4 weeks apart before first documentation of PD

(and not qualifying for a CR). Two PRs separated by one or more SD is considered a confirmed PR as long as the two PR responses are observed ≥4 weeks apart.

- **SD**: At least one objective status of SD or better documented ≥6 weeks after the date of randomization and before first documentation of PD (and not qualifying for CR or PR).

- **Non-CR/non-PD** (applicable only to Phase 2 portion patients with non-measurable disease at baseline): At least one non-CR/non-PD assessment (or better) ≥6 weeks after the date of randomization and before first documentation of PD (and not qualifying for CR or PR).

- **PD**: Documentation of objective status of PD prior to the second planned assessment (ie, ≤12 weeks) after the date of randomization (and not qualifying for CR, PR, SD or non-CR/non-PD).

- **Not Evaluable (NE)**: All other cases.

Reasons for NE will be summarized with the following reasons for each treatment arm:

- No adequate baseline assessment
- No post-baseline assessments due to early death (<6 weeks after treatment start date)
- No post-baseline assessments due to other reasons
- All post-baseline assessments have overall response NE
- New anticancer therapy started before first post-baseline tumor assessment
- SD too early (<6 weeks after the date of randomization without further evaluable tumor assessments)
- PD too late (after two inadequate assessments or after missing two consecutive planned assessments, ie, >12 weeks after the date of randomization)

ORR in the Phase 2 portion will be estimated for each treatment arm as the proportion of patients with a BOR of confirmed CR or PR relative to the FAS and the subset of FAS with measurable disease at baseline, respectively. Corresponding 2-sided 95% CIs will be calculated using the Wilson method. ORR will be compared between the two treatment arms using a 1-sided exact Cochran-Mantel-Haenszel (CMH) test stratified by the randomization stratification factor. The stratified odds ratio and its 2-sided 95% CI will be reported. Unstratified analysis may be used if it is deemed that the strata have too few patients to conduct robust stratified analyses.

6.2.2.3. DOR

DOR is defined, for patients with confirmed OR, as the time from the first documentation of OR to the first documentation of objective PD or to death due to any cause, whichever occurs first. For the Phase 2 portion, DOR will be summarized separately based on investigator assessment and BIRC results.

DOR (in months) is calculated as (date of event or censoring – first date of OR + 1)/30.4375.

Patients last known to be alive and progression-free will be censored at the date of the last objective tumor assessment that verified lack of disease progression, regardless of initiation of new anticancer therapy. In addition,

- Patients with documentation of progression or death after 2 or more missed or inadequate post-baseline assessments will be censored at the time of last objective tumor assessment documenting no progression.

- If patients are removed from the study (eg, withdrew the consent, lost to follow up) prior to an event, DOR will be censored at the date of the last objective tumor assessment that verified lack of disease progression.

DOR will be analyzed for the subgroup of patients with confirmed OR. Kaplan-Meier methods will be used to describe DOR if warranted. The median event time and 2-sided 95% CI for the median will be provided. Reasons for censoring and category hierarchy will follow those for the analysis of EFS; however, patients will not be censored for inadequate baseline assessment or for no adequate post-baseline assessment, as only patients with an objective response will be included in the analysis of DOR.

For dose escalation part and dose determination part of the Phase 1 portion, if there are <3 responses per dose level, no DOR table will be presented. Instead, DOR data (duration of response, treatment/cohort) will be included in efficacy listing.

6.2.2.4. PFS

PFS is defined as the time from the 'start date' (date of first dose of study treatment for the Phase 1 portion and date of randomization for the Phase 2 portion) to the date of the first documentation of objective PD or death due to any cause, whichever occurs first. If tumor progression data include more than one date, the first date will be used. PFS (in months) will be calculated as (date of event or censoring – first dose date +1)/30.4375. PFS for the Phase 2 portion will be summarized separately based on investigator assessment and BIRC results.

Patients last known to be alive and progression-free will be censored at the date of the last objective tumor assessment that verified lack of disease progression, regardless of initiation of new anticancer therapy. In addition,

- Patients with no baseline tumor assessment (including patients with an inadequate baseline assessment) will be censored on the 'start date' with duration of 1 day.

- Patients with no adequate post-baseline tumor assessments will be censored on the 'start date' with duration of 1 day unless death occurred prior to the second planned assessment in which case death will be an event on date of death.

- Patients with documentation of progression or death after 2 or more missing or inadequate post-baseline assessments since the last tumor assessment will be censored at the time of last objective tumor assessment documenting no progression.

- Patients who are removed from the study (withdrawal of consent, lost to follow-up, etc.) prior to progression and death will be censored at the date of the last objective tumor assessment that verified lack of disease progression.

Kaplan-Meier estimates will be presented with a summary of associated statistics including the median PFS time with the 2-sided 95% CI. PFS (%) at specific time points (eg, 3, 6, 12, 18, 24, 36 months) will be estimated with corresponding 2-sided 95% CIs.

Number and percentage of patients with each event type (objective progression, death without objective progression) and censoring reasons will be summarized following the same category hierarchy shown in Table 9.

6.2.2.5. OS

OS is defined as the time from the 'start date' (date of first dose of study treatment for the Phase 1 portion and date of randomization for the Phase 2 portion) to the date of death due to any cause. OS (in months) is calculated as (date of death or censoring – first dose date +1)/30.4375. Patients last known to be alive will be censored at date of last contact.

Kaplan-Meier estimates will be presented with a summary of associated statistics including the median OS time with 2-sided 95% CI. OS (%) at specific time points (eg, 6, 12 and 24 months) will be estimated with corresponding 2-sided 95% CIs.

Number and percentage of patients with an event (death) and censoring reasons will be summarized. Reasons for censoring will be summarized following the category hierarchy shown in Table 10:

Table 10. OS Censoring Reasons and Hierarchy

Hierarchy	Condition	Censoring Reason
1	No event and withdrawal of consent date ≥ date of first dose of study treatment (Phase 1 portion) or date of randomization (Phase 2 portion)	Withdrawal of consent
2	No event and lost to follow-up in any disposition or survival follow-up page	Lost to follow-up
3	No event and none of the conditions in the prior hierarchy are met	Alive

Follow-up duration for OS will be described using reverse Kaplan-Meier method (reversing the OS censoring and event indicators). Associated statistics including the median time of follow-up for OS with 2-sided 95% CIs will be provided.

6.2.2.6. PET-CT/MRI response after Cycle 4 (Phase 2 portion only)

FDG-PET or PET-CT/MRI, if performed at baseline for distant metastases, will be repeated after Cycle 4 (at completion of Cycle 4; not mandatory). If increased FDG uptake alone on FDG-PET corresponds to a target lesion that has enlarged, but the sum of the target measurements does not show an increase that is sufficient for PD (other target lesions have not enlarged or have actually decreased), the FDG-PET does not override the assessment following RECIST 1.1.

This analysis will be performed in the subset of the FAS with FDG-PET response and MRI/CT objective response assessed at baseline and after Cycle 4. FDG-PET response after Cycle 4 will be descriptively compared to objective response on MRI/CT. Lesion level of FDG-PET uptake intensity or avidity will be recorded as Resolved, Improved, Stable, Increased, First appearance, NE. The number and percentage of patients with FDG-PET overall response at Cycle 4 (derived) will be tabulated against the MRI/CT objective response status (CR, PR, SD, PD, NE) at Cycle 4. Corresponding lesion level FDG-PET response (ie, uptake intensity compared to baseline) and MRI/CT assessments at Cycle 4 will also be presented in a patient listing.

6.2.3. PK Endpoints

The PK analyses will be based on the PK analyses set.

6.2.3.1. PK Analysis of Palbociclib, IRN, TMZ, TOPO and CTX

Plasma PK parameters including the maximum plasma concentration (C_{max}), time to maximum plasma concentration (T_{max}), and predose concentration (C_{trough}) will be estimated for all measured analytes using noncompartmental analysis as data permit. If data permit or if considered appropriate, area under the plasma concentration vs. time curve ($AUC\tau$), apparent oral plasma clearance (CL/F) or CL for IV dosing will also be estimated.

Palbociclib steady-state PK parameters will be summarized descriptively by dose level (mg/m^2) separately for each combination (n, mean, standard deviation, %CV, median, minimum, maximum, geometric mean and its associated %CV) as appropriate per parameter. Box and whisker plots for individual subject exposure parameters (AUC_t, C_{max} and C_{trough}) will be presented by dose-level and combination.

Dose-normalized palbociclib PK parameters (normalized by mg/m^2 dose level) will be pooled including all evaluable parameters from the Phase 1 and Phase 2 portions and summarized by prespecified age groups (≤6 years, >6 years and <12 years, < 12 years, and ≥12 years and <18 years; ≥18 years, and <18 years). This summary will be created separately for each combination and additionally for pooled palbociclib PK parameters across the two combinations. Box and whisker plots for individual subject dose-level normalized exposure parameters (AUC_t, C_{max} and C_{trough}) will be presented by selected age groups and overlaid with individual values and geometric means.

Steady-state PK parameters of palbociclib combination partners (IRN, SN-38, TMZ, TOPO and CTX) will be summarized descriptively (n, mean, standard deviation, %CV, median,

minimum, maximum, geometric mean, and its associated %CV) as appropriate per parameter.

Analyte concentrations will be summarized descriptively separately for each combination by analyte, dose, cycle, day, and nominal time (n, mean, standard deviation, %CV, median, minimum, maximum, geometric mean, and its associated %CV). Individual participant (spaghetti plots), as well as median profiles of the plasma concentration-time data for each analyte will be plotted by each combination using actual (for individual) and nominal (for median) times respectively.These plots may be further stratified by prespecified age groups as mentioned above, Palbociclib dose level or be presented in a dose-normalized manner.

Dose-normalized palbociclib PK concentrations (normalized by mg/m^2 dose level) will be pooled including all evaluable concentrations from the Phase 1 and Phase 2 portions and will be summarized by prespecified age groups (≤6 years, >6 years and <12 years, <12 years, and ≥12 years and <18 years; ≥18 years, and <18 years). This summary will be created separately for each combination and additionally for pooled palbociclib PK concentrations across the two combinations. Dose-normalized palbociclib concentration-time plots from serial PK collection visits will be generated on both linear and semi-log scales and will be presented by overlaying selected age groups.

The relationship between analyte concentrations and/or PK parameters with response, biomarkers and/or safety findings after accounting for potential covariates may be evaluated. The results of any exploratory correlational and/or population PK/PD analysis may be reported separately from the clinical study report.

6.2.3.2. Population PK Analysis or PK/PD Modeling

PK and PD data from this study may be analyzed using modeling approaches to investigate any association between palbociclib exposure and PD endpoints (eg, TK1 activity) or significant safety and/or efficacy endpoints. The results of these analyses, if performed, may be combined with data from other studies, as appropriate, and may be reported separately.

6.3. Other Endpoints

6.3.1. PRO Endpoints (Phase 2 portion)

The following analyses will be performed by treatment arm and age appropriate questionnaire based on the FAS. All PRO analyses will be considered descriptive and exploratory in this study.

6.3.1.1. PRO questionnaires

PROs will be collected at baseline, after completion of Cycle 2 study treatment and preferably immediately prior to Cycle 3 Day 1 (pre-dose) and after completion of Cycle 4 study treatment and preferably immediately prior to Cycle 5 Day 1 (pre-dose) using the following age appropriate questionnaires:

- PROMIS Early Childhood Parent Report Scale v1.0 - Global Health 8a (age ≥2 to <5 years)

- PROMIS Parent Proxy Profile v2.0 – Profile-25 (age ≥5 to ≤17 years)
- PROMIS Pediatric Profile v2.0 (age ≥ 8 to ≤ 17 years)
- PROMIS–29 Profile v2.1 (age ≥18 years)

6.3.1.2. PRO scoring

The PROMIS questionnaires will be scored according to instrument developer guidance.

6.3.1.3. PRO questionnaire completion status

For each treatment arm and at each time point, completion status of each questionnaire (ie, completed all questions, completed at least one question, and completed no questions) will be summarized with frequencies and percentages by questionnaire, including the reasons for non-completion.

6.3.1.4. Descriptive summary

Descriptive statistics (mean, SD, median, range and 95% CI) will be used to summarize, for each treatment arm and at each time point, the absolute scores and change from baseline by questionnaire defined by age groups (Section 6.3.1.1) and corresponding domain or scale.

6.3.2. Days of hospitalization (Phase 2 portion)

Average and maximum days of hospitalization while on study will be calculated for each patient and descriptive statistics (mean, SD, median, minimum, and maximum) will be presented for each treatment arm, respectively. Admittance reasons will be summarized by treatment arm according to categories collected on the Hospitalization Details CRF page. The reporting period for hospitalization starts with Cycle 1 Day 1 and continues through out study treatment period until the EOT visit.

6.3.3. Exploratory Biomarker Endpoints

Biomarker data will be analyzed based on the biomarker analysis sets. The extent of analyses of biomarker endpoints will depend on availability of samples collected from pediatric patients in this study.

Biomarkers will be assessed separately for blood samples and archival tumor tissue. In each case, summary statistics (eg, mean, SD, median, minimum, maximum levels of continuous, and frequencies and percentages of categorical biomarker measures) will be determined at baseline and post-treatment change from baseline (where appropriate).

Patients may be classified as positive, negative, or some other category according to scoring algorithms and cutoffs established from external sources. If no external standards exist, patients may be stratified using the median, quartiles and tertiles. The number and percentage of patients in each category will be tabulated. The number of OR will be tabulated relative to biomarker classifications using a contingency table as appropriate.

Data from biomarker assays may be analyzed using graphical presentation and statistical methods such as, but not limited to linear regression, t-test, and analysis of variance to explore any possible relationship of biomarker levels and status and/or changes from baseline with clinical outcomes. The statistical approach will also examine correlations of biomarker results with PK parameters and measures of antitumor efficacy.

The exploratory biomarker endpoints may be reported only if data allow and may not all be included in the CSR. Results of exploratory research may be shared through presentation(s) to conferences and/or manuscript(s).

6.3.4. Palatability of Palbociclib Oral Formulation

Taste and palatability of the palbociclib oral formulation will be assessed at the time of the first dose, Cycle 1 Day 14, Cycle 2 Day 1, and Cycle 2 Day 14 only for patients taking the palbociclib oral solution orally using the Palbociclib Oral Solution Taste Feedback Questionnaire (Appendix 9). At each time point the number and percentage of patients who complete the palatability questionnaire will be summarized, as will the reasons for non-completion of the measures. A questionnaire is considered complete if at least one item was answered by the patient/caregiver.

For the question “How much do you like this medicine overall?”, palatability is assessed on a scale of 1 to 5 where 1=dislike very much, 2=dislike a little, 3=not sure, 4=like a little, and 5=like very much. Descriptive statistics (mean, SD, median, minimum, and maximum) of the observed taste score and change from baseline may be provided at each time point, if warranted. Number and percentage of patients in each palatability category will be provided at each time point. Summaries of successful dosing (yes/no) and trouble swallowing the medicine if dosing was successful will also be provided.

6.3.5. Bone Metabolism and Growth

Assessments of growth and bone metabolism will be performed in a subset of patients in the safety analysis set who have Tanner stage less than Tanner 4 at the time points described in the Schedule of Activities per protocol. Parameters (eg, linear growth, bone mineral density, bone age, Tanner stage, insulin-like growth factor 1 [IGF]), luteinizing hormone [LH], follicle stimulating hormone [FSH], estradiol for girls, testosterone for boys, bone alkaline phosphatase and C-telopeptide cross-links [Phase 2 portion only]) will be summarized using descriptive statistics (mean, SD, median, minimum, and maximum) of actual values at baseline and maximum change from baseline over time. Number and percentage of patients with bone age by X-ray less than, equal or greater than their calendar age will be summarized at baseline and each time point. Descriptive summary (mean, SD, median, minimum, and maximum) of difference between bone age and calendar age will also be provided.

6.4. Subset Analyses

Not applicable for the Phase 1 portion due to the small sample size and exploratory nature of the study.

For the Phase 2 portion, subset analyses will be performed for investigator-assessed EFS based on the FAS for the subgroups below and other baseline patient demographics and

disease characteristics described in Section 6.5.1, if warranted. All the subset analyses are exploratory. No adjustment for multiplicity will be performed.

- Randomization stratification factor
  - Type and time of current disease recurrence (primary refractory or 1st recurrence <2 years, 1st recurrence ≥2 years since completion of first line treatment or 2nd or greater recurrence).

- Age (age ≥2 and <18 years, ≥18 years)

- Prior treatment combination of IRN and/or TMZ (Yes, No)

Subset analyses for EFS will use the censoring rules described in Section 6.1.4.1. The unstratified HR and its corresponding 95% CI will be provided per subgroup level. In the case of a low number of patients within a category (<10% of the randomized population), the categories will be pooled, or in cases where no meaningful pooling can be performed, the category may not be summarized.

6.5. Baseline and Other Summaries and Analyses

6.5.1. Baseline Summaries

Unless otherwise specified, the summaries (eg, number and percentage of patients for binary/categorical variables or summary statistics for continuous variables) will be provided based on the FAS (safety analysis set) overall and by dose level for the Phase 1 portion and by treatment arm for the Phase 2 portion.

- Demographic characteristics
  - Age (years): summary statistics
  - Age categories: (≤6, >6 - <12, ≥12 - <18, and ≥18 years; <18 and ≥18 years)
  - Gender: Male, Female
  - Race: White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Not Reported
  - Ethnic origin: Hispanic or Latino(a) or of Spanish Origin, Not Hispanic or Latino(a) or of Spanish Origin, Not Reported
  - Geographic region (Phase 2 portion only, as applicable): North America, Europe and Israel, ROW

- Physical measurements
  - ECOG (>16 years): 0, 1, 2
  - Lansky Performance Status (≤16 years): 90-100%, 70-80%, 50-60%
  - Tanner stage: Stage 1-5 (girls younger than 8 years old or boys younger than 9 years old will be considered as Tanner stage of 1 unless otherwise specified)

- Randomization stratification according to IRT
  - Type and time of current disease recurrence: primary refractory or 1st recurrence <2 years, 1st recurrence ≥2 years since completion of first line treatmentor 2nd or greater recurrence.

The stratification factor will be summarized by treatment, as well as the concordance between the data recorded in the IRT at randomization and the eCRF by treatment.

- Medical history
  - Medical history will be coded using the latest version of MedDRA and will be summarized as the numbers and percentages of patients by MedDRA preferred term (PT) as event category and MedDRA primary system organ class (SOC) as summary category. Each patient will be counted only once within each PT or SOC. Medical history will be displayed in terms of frequency tables: ordered by primary SOC and PT in alphabetical order.

- Disease characteristics
  - Primary diagnosis (histopathological classification)
  - Duration since primary diagnosis (histopathological diagnosis) to the 'start' date (years): summary statistics
  - Molecular diagnosis: EWS, EWS-like, Other, Unknown (Phase 2 portion only)
  - Site of recurrence: Locally advanced, Metastatic (Phase 2 portion only)
  - Measurable disease: Yes, No (Phase 2 portion only)

- Prior anticancer therapies

  The number and percentage of patients in each of the following anticancer therapy categories will be tabulated:

  - Patients with at least one prior anticancer therapy
  - Prior anticancer drug therapy (Yes, No), and number of regimens (1, 2, 3, ≥4) if prior anticancer drug therapy was received.
  - Prior anticancer radiation therapy (Yes, No).
  - Prior anticancer surgery (Yes, No)

6.5.2. Study Conduct and Patient Disposition

6.5.2.1. Patient Evaluation Groups

Phase 1 Portion and Tumor-Specific Cohort(s)

- Number of patients screened
- Number of patients with screen failures
- Number of patients not screen failures but not enrolled
- Number of patients enrolled
  - Number and percentage of patients treated
  - Number and percentage of patients not treated
- Number and percentage of treated patients in each of the analysis sets defined in Section 4

Randomized Phase 2 Portion

- Number of patients screened
- Number of patients with screen failures
- Number of patients not screen failures but not randomized
- Number of patients randomized
    - Number and percentage of patients randomized but not treated
    - Number and percentage of patients treated
- Number and percentage of treated patients in each of the analysis sets defined in Section 4

6.5.2.2. Patient Disposition

- Number and percentage of patients who entered treatment phase
- Number and percentage of patients who discontinued study treatment and primary reason for discontinuation
- Number and percentage of patients with study treatment ongoing
- Number and percentage of patients who entered follow-up phase
- Number and percentage of patients who discontinued follow-up overall and primary reason for discontinuation
- Number and percentage of patients ongoing in follow-up

6.5.3. Protocol Deviations

Potentially important protocol deviations (identification and categorization follows the Pfizer standard process) that impact the safety of the patients and/or the conduct of the study and/or its evaluation will be summarized, which may include but are not limited to, eligibility violations, deviations from the conduct of the study after start of study treatment, or cases where important study evaluations were not performed as required by the protocol.

6.5.4. Impact of COVID-19

The potential impact of COVID-19 on the study will be reported as the following, as applicable.

- Patient discontinuations related to COVID-19 will be summarized.

- A listing of all patients affected by COVID-19 related study disruption by unique subject number identifier and by investigational site, and a description of how the individual's participation was altered (if data collection supports).

- A listing of Protocol Deviations related to COVID-19. A listing of protocol deviations related to COVID-19 will be provided. The identification of these and other CSR-reportable deviations will be based on the inclusion/exclusion criteria or other criteria presented in the protocol.

6.5.5. Study Treatment Exposure and Compliance

Details of treatment exposure are described in Appendix 5.

The following analyses will be summarized based on the safety analysis set by individual study drug.

6.5.5.1. Exposure to study drug

The exposure to each study drug will be summarized using descriptive statistics [mean (SD), median (minimum, maximum)] for all cycles combined:

- Duration of exposure to palbociclib
  - Duration of treatment (months) = (last dose date – first dose date + d) / 30.4375, where d = 1 if the patient discontinues study treatment prior to the end of last cycle (ie, if date of last dose of study treatment – date of first dose of study treatment in the last cycle < 14) or d = 1 + 7 otherwise, regardless of dose interruptions or cycle delay.
- Duration of exposure to IRN, TMZ, TOPO or CTX
  - Duration of treatment (months) = (last dose date – first dose date + d) / 30.4375, where d = 1 if the patient discontinues study treatment prior to the end of last cycle (ie, if date of last dose of study treatment – date of first dose of study treatment in the last cycle < 5) or d = 1 + 16 otherwise, regardless of dose interruptions or cycle delay.
- Number of dosed cycles
  - For all cycles including the last cycle in which doses were given even if the cycle was not completed.
- Cumulative dose
  - Cumulative dose (mg) = sum of all actual doses the patient received from Cycle 1 to and including the last cycle.
- Actual total dose intensity
  - Actual total dose intensity (mg/day) = actual total dose taken over the entire treatment duration divided by actual duration of treatment including delays
- Intended total dose intensity
  - Intended total dose intensity (mg/day) = intended total dose over the entire treatment duration divided by intended duration of treatment
- Relative total dose intensity
  - Relative total dose intensity (%) = actual total dose intensity divided by the intended total dose intensity

6.5.5.2. Treatment Modifications

Dose modifications of study treatments will be summarized as follows including numbers and percentages:

- Patients with at least one dose reduction at any time during drug administration, patients with at least one dose reduction due to AE, and a breakdown of the number of dose reductions (eg, 1, 2, 3). Dose reduction is defined as a change to a non-zero dose level lower than the prior dose level (planned dose).

- Patients with at least one dose interruption at any time during drug administration, patients with at least one dose interruption due to AE, a breakdown of the number of dose interruptions (eg, 1, 2, 3, ≥4), and duration of interruption (maximum and average). Dose interruption is defined as a planned dosing day with 0 mg administered within a cycle.

- Patients with at least one cycle delay (ie, start of following cycle is delayed) at any time during drug administration, a breakdown of the number of cycle delays (eg, 1, 2, 3, ≥4), and duration of cycle delay (maximum and average; palbociclib only). A cycle was deemed to have been delayed if [start date of Cycle X – start date of Cycle (X-1) – 21] > 3.

6.5.6. Concomitant Medications and Non-Drug Treatments

Concomitant medications and non-drug treatments refer to all drug and non-drug treatments, other than study medications, which started prior to first dose date of study treatment and continued during the on-treatment period as well as those started during the on-treatment period. Concomitant medication will be summarized based on the safety analysis set as shown in Table 7 and Table 8.

Summary of concomitant medications will include the number and percentage of patients by Anatomical Therapeutic Chemical (ATC) Classification level 2 and preferred term. A patient will be counted only once within a given drug class and within a given drug name, even if he/she received the same medication at different times. If any concomitant medication is classified into multiple ATC classes, the medication will be summarized separately under each of these ATC classes. The summary tables will be sorted on decreasing frequency of drug class and decreasing frequency of drug name in a given drug class. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used. In case any specific medication does not have ATC classification level 2 coded term, it will be summarized under 'Unavailable ATC classification' category.

6.5.7. Follow-Up Cancer Therapies

Follow-up cancer therapies after treatment discontinuation will be summarized based on the FAS. Number and percentage of patients with any follow-up cancer therapy will be tabulated overall and by type of therapy (Follow-up Cancer Therapy, Follow-up Radiation Therapy, Follow-up Cancer Surgery).

6.6. Safety Summaries and Analyses

The safety analysis set will be the primary population for safety evaluations.

For the Phase 1 portion, AEs and other safety parameters will be summarized for each treatment combination overall and by dose level. For tumor-specific cohorts, AEs and other safety parameters will be summarized by tumor type. For the Phase 2 portion, AEs and other safety parameters will be summarized by treatment arm.

6.6.1. Adverse Events

TEAEs occurring during the on-treatment period are defined as in Section 3.5.1.1. All analyses described will be based on TEAEs if not otherwise specified. AEs outside the on-treatment period will be flagged in the listings.

Unless otherwise specified, AEs will be summarized with number and percentage of patients by MedDRA SOC and PT in decreasing frequency. In addition, AEs will also be summarized by severity and relationship to study drug by all cycles and repeated for Cycle 1 for the Phase 1 portion only.

Each patient will be counted only once within each SOC or PT. If a patient experiences more than one AE within a SOC or PT for the same summary period, only the AE with the strongest relationship or the worst severity, as appropriate, will be included in the summaries of relationship and severity. If relationship to study drug is missing, the AE will be counted as treatment related. AE listings will show missing relationship as missing.

In case a patient has events with missing and non-missing grades, the maximum of the non-missing grades will be displayed. No imputation of missing grades will be performed.

The denominator for the percentage of patients with an event will be based on the number of patients in the safety analysis set unless otherwise specified.

6.6.1.1. All Adverse Events and Serious Adverse Events

The number and percentage of patients with each of the following will be presented by treatment combination overall and by dose level (Phase 1 portion), by tumor type (tumor-specific cohorts) and by treatment arm (Phase 2 portion). For the Phase 2 portion of the study, all SAEs, Grade ≥3 non-serious AEs, and any AEs leading to study treatment interruptions, requiring dose reduction or leading to permanent discontinuation of study treatment are collected.

- Overall summary of TEAEs will include the number and percentage of participants with the following:
    - TEAEs
    - Grade 3 or 4 TEAEs
    - Grade 5 TEAEs
    - Treatment related TEAEs
    - Grade 3 or 4 treatment related TEAEs
    - Grade 5 treatment related TEAEs
    - Treatment-emergent SAEs
    - Treatment related treatment-emergent SAEs
    - TEAEs leading to dose interruption of any study drug
    - TEAEs leading to dose interruption of palbociclib
    - TEAEs leading to dose reduction of any study drug
    - TEAEs leading to dose reduction of palbociclib
    - TEAEs leading to permanent discontinuation of all study drugs

  - Treatment related TEAEs leading to permanent discontinuation of all study drugs

- TEAEs by MedDRA SOC, PT and maximum CTCAE Grade (All Causality, Treatment Related, All Cycles)

- TEAEs by MedDRA PT (including clusters of PTs) and maximum CTCAE Grade (All Causality, Treatment Related, All Cycles)

- Treatment-emergent SAEs by MedDRA SOC and PT (All Causality, Treatment Related, All Cycles)
- Treatment-emergent SAEs by MedDRA PT (including clusters of PTs) and maximum CTCAE Grade (All Causalities, All Cycles)

- Treatment-emergent SAEs by MedDRA PT (including clusters of PTs) (All Causalities, All Cycles)

- Duration of Grade 4 NEUTROPENIA and THROMBOCYTOPENIA by MedDRA PT (All Cycles)

Number and percentage of patients with COVID-19 Related TEAEs by SOC and PT, as noted on the eCRF as an AE with causality " and specified as "Unconfirmed COVID-19 infection" or "Positive SARS-CoV-2 test", will be summarized as applicable.

6.6.1.2. Adverse Events Leading to Dose Modification

The number and percentage of patients with each of the following will be presented for TEAEs leading to dose modification (dose reduction, dose interruption) by treatment combination overall and by dose level (Phase 1 portion), by tumor type (tumor-specific cohorts) and by treatment arm (Phase 2 portion):

- TEAEs leading to dose reduction of any study drug by MedDRA PT (including clusters of PTs) (All Causality, Treatment Related, All Cycles)

- TEAEs leading to dose reduction of any study drug by MedDRA PT (including clusters of PTs) (All Causality, Treatment Related, All Cycles)

- TEAEs leading to dose reduction of palbociclib by MedDRA PT (including clusters of PTs) (All Cycles)

- TEAEs leading to dose interruption of any study drug by MedDRA PT (including clusters of PTs) (All Causality, Treatment Related, All Cycles)

- TEAEs leading to dose interruption of palbociclib by MedDRA PT (including clusters of PTs) (All Cycles)

6.6.1.3. Adverse Events Leading to Permanent Discontinuation of Study Treatment

- TEAEs leading to permanent discontinuation of study drugs by MedDRA PT (including clusters of PTs) (All Causality, Treatment Related, All Cycles)

6.6.1.4. Adverse Events Leading to Death

- TEAEs leading to death by MedDRA PT (All Causality, Treatment Related, All Cycles)

6.6.2. Deaths

Summary of death will be presented with numbers and percentages as following:

- All deaths
- Deaths from start of treatment to and including 35 days after last dose of study treatment
- Deaths during follow-up period occurring greater than 35 days after last dose of study treatment
- Cause of Death
    - Disease under study
    - Study treatment toxicity
    - COVID-19 (if any)
    - Other
    - Unknown

In addition, date of death, cause of death, and flag for death within 35 days of last dose of study treatment will be provided in individual patient listing together with selected dosing information (date of first/last dose, dose) and AEs with fatal outcome (list PT of AEs with outcome=Fatal, and AEs of Grade 5).

6.6.2. Laboratory Data

Laboratory parameters will be summarized for all cycles and Cycle 1 (Phase 1 portion only), respectively. Hematology and clinical chemistry laboratory results will be classified according to the NCI CTCAE v.4.03.

Laboratory parameters that have criteria available for both low and high values will be summarized for both criteria (low and high). Patients will only be counted once for each criterion. The same patient can be counted for both criteria if the patient has laboratory values meeting each criterion.

The number and percentage of patients who experienced laboratory test abnormalities will be summarized according to worst toxicity grade observed for each lab parameter during the on-treatment period. The denominator to calculate percentages for each laboratory parameter is the number of patients evaluable for CTCAE grading (ie, those patients for whom a Grade 0, 1, 2, 3 or 4 can be derived).

Other laboratory parameters such as coagulation (prothrombin time (PT) or international normalized ratio (INR) and pregnancy test collected on the eCRF will be listed.

The listings of laboratory results will be provided for all laboratory parameters including CTCAE grading. The listings will be sorted by parameters and assessment dates or visits for each patient. Laboratory values that are outside the normal range will also be flagged in the data listings, along with corresponding normal ranges.

For the Phase 2 portion, the following additional analyses will be summarized:

- Shift summary of hematology laboratory test results during the on-treatment period by maximum CTCAE grade
- Shift summary of hematology laboratory test results with no CTCAE criteria by worst on-treatment assessment
- Shift summary of chemistry laboratory test results during the on-treatment period by maximum CTCAE grade
- Shift summary of chemistry laboratory test results with no CTCAE criteria by worst on-treatment assessment

6.6.3. Vital Signs and Physical Examination

Vital sign summaries will be performed for the Phase 1 portion and tumor-specific cohort(s) only, including all vital sign assessments at baseline and during the on-treatment period.

All vital sign quantitative parameters (eg, weight, height, BMI, BSA, blood pressure, pulse rate) will be summarized using descriptive statistics (mean, SD, median, minimum, and maximum) of actual values for baseline and all cycles with maximum change from baseline. The changes computed will be the differences from baseline. Unscheduled measurements will not be used in these summaries.

6.6.4. ECG

All ECGs will be recorded in triplicate, ie, three ECGs taken approximately 2 minutes apart. ECG parameters (PR, QRS, QT, QTcF, and QTcB). After the above parameters have been derived within each patient and scheduled time point, each ECG parameter should be averaged as follows: (1^{st} measurement + 2^{nd} measurement + 3^{rd} measurement)/3. Any data obtained from ECGs repeated for safety reasons after the nominal timepoints will not be averaged along with the preceding triplicates. QTcF (QT interval corrected using Fridericia's formula) and QTcB (QT interval corrected using Bazett's formula) will be derived based on heart rate (HR) (or RR) and QT. The mean value of the triplicate measurements should be determined after the derivation of the individual parameter at each time point.

Change from baseline calculations for ECG measurements will be derived from the triplicate averaged measurements. The baseline ECG assessment is defined as the triplicate ECG assessment taken pre-dose on Cycle 1 Day 1, or the most recent triplicate ECG assessment reported prior to the first administration of study drug. Change from baseline is defined as a

patient's parameter value at a particular time point minus the appropriately matched baseline value. Change from baseline calculations should only use post-dose ECG measurements.

ECG parameters during the on-treatment period (including maximum post-baseline values) will be summarized with the following categories. All post-dose QTc interval data should be used in determining the maximum for a patient, including all scheduled and unscheduled ECG results.

- QTc ≤ 450 msec
- QTc > 450 msec but ≤ 480 msec
- QTc > 480 msec but ≤ 500 msec
- QTc > 500 msec

ECG parameters during the on-treatment period (including maximum change from baseline values) will be summarized with the following categories:

- Change ≤ 30 msec
- Change > 30 msec but ≤ 60 msec
- Change > 60 msec

All ECG assessments will be listed, and those collected outside the on-treatment period will be flagged in the listing. Patients with qualitative ECG abnormalities will be listed for each patient and time point and the corresponding abnormality findings will be included in the listings.

7. INTERIM ANALYSES

7.1. Introduction

Phase 1 Portion

No formal interim analysis will be conducted for the dose escalation part, dose determination part, and dose expansion cohort of the study. However, as this is an open-label study, the sponsor may conduct unblinded reviews of the data during the course of the study for the purpose of safety assessment, facilitating dose-escalation/dose-determination decisions, facilitating PK/PD modeling, and/or to support clinical development.

Tumor-Specific Cohort(s)

For the tumor-specific cohorts (if opened) with a 2-stage design, there will be 1 formal interim analysis for futility and efficacy after the first stage of accrual is complete and the last patient was followed for 6 months after the first dose of study treatment. Unless criteria for continuation have already been met, accrual will be temporarily suspended for assessment of the primary endpoint after the first stage of accrual. If continuation criteria are not met or pre-defined futility/efficacy criteria is met (Table 6) after assessments are complete, the corresponding tumor-specific cohort will be permanently closed.

Randomized Phase 2 Portion

The Phase 2 portion of the study is designed to have one interim futility analysis and the final analysis for the primary endpoint of investigator-assessed EFS based on the FAS. The goal of the interim analysis is to allow for early stopping of the study due to futility/no signal of activity. A group sequential design with a Rho family (1.0) β-spending function is used to determine the futility boundary (binding). The interim analysis of EFS will be performed as described in Sections 5.1.1 and 5.1.2 using the methodology described in Section 6.1.4.

Unblinded results from the interim analysis for EFS will not be communicated to the Sponsor's clinical team or to any party involved in the study conduct (apart from the independent reporting statistician and E-DMC members) until the E-DMC has determined that the primary EFS analysis has crossed the pre-specified boundary for futility or the study needs to be terminated due to any cause, including safety reasons. Further details are described in the E-DMC charter.

7.2. Interim Analyses and Summaries

Tumor-Specific Cohort(s)

Interim analysis (Stage 1) for the primary efficacy endpoint (ie, OR) within the tumor-specific cohort, if opened, by treatment combination will be summarized as described in Section 6.1.2.

Patients with the specific tumor type treated in the dose escalation part (for combination of palbociclib with IRN and TMZ) or dose determination part (for combination of palbociclib with TOPO and CTX) who showed response at any dose level and those with the specific tumor treated at the same dose level who did not respond, or treated at the MTD or the potential RP2D regardless of response as well as patients in the respective dose expansion cohort will be counted towards the number of patients for Stage 1.

Randomized Phase 2 Portion

The interim analysis for futility will be performed based on investigator-assessed EFS after approximately 30 EFS events (60% of information for EFS) have been documented. If the 1-sided p-value is >0.4287 ($Z > -0.1796$) at the interim analysis, the Phase 2 portion of the study will be terminated early for lack of efficacy. The futility boundary will be updated where applicable based on the acutal number of observed EFS events for the interim analysis using the pre-specified β-spending function as described in Sections 5.1.2.

Final analysis of EFS will be performed after approximately 50 EFS events (100% of information for EFS) have been documented. The study would be considered to meet its primary endpoint if the 1-sided log-rank test (stratified) for EFS based on the FAS is significant at the significance level of 0.20 at the final analysis (with ~50 events).

8. REFERENCES

1. Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. American Statistician 1998; 52:119–126.

2. Martin GM, et al. Results of the second interim assessment of rEECur, an international randomized controlled trial of chemotherapy for the treatment of recurrent and primary refractory Ewing sarcoma (RR-ES). 2020 ASCO Annual Meeting. Abstract 11502: https://meetinglibrary.asco.org/record/185567/abstract

3. Brookmeyer R, Crowley JJ. A confidence interval for the median survival time. Biometrics 1982; 38:29-41.

4. Cox DR. Regression models and life-tables (with discussion). Journal of the Royal Statistical Society, Series B 1972; 34:187-220.

5. Eisenhauer E, Therasse P, Bogaerts J, Schwartz LH, et al. New response evaluation criteria in solid tumors: Revised RECIST guideline version 1.1. Eur J Can 45: 228-47, 2009.

6. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958; 53:457-481.

7. ICH Harmonised Tripartite Guideline. The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs. E14. October 2005.

8. Mander AP, Thompson SG. Two-stage designs optimal under the alternative hypothesis for phase II cancer clinical trials. Contemp Clin Trials. 2010; 31(6): 572-8.

9. McCabe MG, Kirton L, Khan M, Fenwick N, Dirksen U, Gaspar N, et al. Results of the second interim assessment of rEECur, an international randomized controlled trial of chemotherapy for the treatment of recurrent and primary refractory Ewing sarcoma (RR-ES). J Clin Oncol. 2020; 38(15_suppl):11502.

10. Park JR, Bagatell R, Cohn SL, et al. Revisions to the International Neuroblastoma Response Criteria: a consensus statement from the National Cancer Institute clinical trials planning meeting. J Clin Oncol. 2017; 35(22): 2580-7.

11. Schoenfeld D. Chi-squared goodness-of-fit tests for the proportional hazards regression model. Biometrika 1980; 67:145-53.

12. Skolnik JM, Barrett JS, Jayaraman B, et al. Shortening the timeline of pediatric phase I trials: the rolling six design. J Clin Oncol. 2008 Jan 10; 26(2):190-5.

13. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. Statistics for Biology and Health 2000.

14. Warren KE, Vezina G, Poussaint TY, et al. Response assessment in medulloblastoma and leptomeningeal seeding tumors: recommendations from the Response Assessment in Pediatric Neuro-Oncology committee. Neuro Oncol. 2018 Jan 10;20(1):13-23.

15. Wen PY, Chang SM, Van den Bent MJ, et al. Response Assessment in Neuro-oncology Clinical Trial. J Clin Oncology 2017 Jul 20; 35 (21): 2439-49.

16. Wilson, EB. Probable inference, the law of succession, and statistical inference. Journal of the American Statistical Association 1927; 22: 209–212.

17. Zhang X. Comparison of restricted mean survival times between treatments based on a stratified Cox model. DOI 10.1515/bams-2013-0101 Bio-Algorithms and Med-Systems 2013; 9(4): 183–9.

9. APPENDICES

Appendix 1. RECIST (Response Evaluation Criteria in Solid Tumors) version 1.1 Guidelines

Adapted from *E.A. Eisenhauer, et al: New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). European Journal of Cancer 45 (2009) 228–247.*[5]

CATEGORIZING LESIONS AT BASELINE

Measurable Lesions

Lesions that can be accurately measured in at least one dimension.

- Lesions with longest diameter twice the slice thickness and at least 10 mm or greater when assessed by CT or MRI (slice thickness 5-8 mm)
- Lesions with longest diameter at least 20 mm when assessed by Chest X-ray
- Superficial lesions with longest diameter 10 mm or greater when assessed by caliper
- Malignant lymph nodes with the short axis 15 mm or greater when assessed by CT.

  NOTE: The shortest axis is used as the diameter for malignant lymph nodes, longest axis for all other measurable lesions.

Non-measurable disease

Non-measurable disease includes lesions too small to be considered measurable (including nodes with short axis between 10 and 14.9 mm) and truly non-measurable disease such as pleural or pericardial effusions, ascites, inflammatory breast disease, leptomeningeal disease, lymphangitic involvement of skin or lung, clinical lesions that cannot be accurately measured with calipers, abdominal masses identified by physical examination that are not measurable by reproducible imaging techniques.

- Bone disease: Bone disease is non-measurable with the exception of soft tissue components that can be evaluated by CT or MRI and meet the definition of measurability at baseline.
- Previous local treatment: A previously irradiated lesion (or lesion subjected to other local treatment) is non-measurable unless it has progressed since completion of treatment.

Normal sites

- Cystic lesions: Simple cysts should not be considered as malignant lesions and should not be recorded either as target or non-target disease. Cystic lesions thought to represent cystic metastases can be measurable lesions, if they meet the specific

definition above. If non-cystic lesions are also present, these are preferred as target lesions.

- Normal nodes: Nodes with short axis <10 mm are considered normal and should not be recorded or followed either as measurable or non-measurable disease.

RECORDING TUMOR ASSESSMENTS

All sites of disease must be assessed at baseline. Baseline assessments should be done as close as possible prior to study start. All disease must be documented appropriately. If baseline assessment is inadequate, subsequent statuses generally should be indeterminate.

Target lesions

- All measurable lesions up to a maximum of 2 lesions per organ, 5 lesions in total, representative of all involved organs, should be identified as target lesions at baseline. Target lesions should be selected on the basis of size (longest lesions) and suitability for accurate repeated measurements. Record the longest diameter for each lesion, except in the case of pathological lymph nodes for which the short axis should be recorded. The sum of the diameters (longest for non-nodal lesions, short axis for nodal lesions) for all target lesions at baseline will be the basis for comparison to assessments performed post-baseline.

- If 2 target lesions coalesce the measurement of the coalesced mass is used. If a large target lesion splits, the sum of the parts is used.

- Measurements for target lesions that become small should continue to be recorded. If the lesion is considered to have disappeared, 0 mm should be recorded; otherwise if a lesion is determined to be present but too small to measure, the lesion status will indicate “too small to measure and judged to be less than 10 mm” and 5 mm will be used in the calculation of the sum of the diameters.

  NOTE: When nodal lesions decrease to <10 mm (normal), the actual measurement should still be recorded.

Non-target disease

All non-measurable disease is non-target. All measurable lesions not identified as target lesions are also included as non-target disease. Measurements are not required but rather assessments will be expressed as ABSENT, INDETERMINATE, PRESENT/NOT INCREASED, INCREASED. Multiple non-target lesions in one organ may be recorded as a single item on the case report form (eg, ‘multiple enlarged pelvic lymph nodes’ or ‘multiple liver metastases’).

Objective Response Status at Each Evaluation

Disease sites must be assessed using the same technique as baseline, including consistent administration of contrast and timing of scanning. If a change needs to be made the case must be discussed with the radiologist to determine if substitution is possible. If not, subsequent objective statuses are indeterminate.

Target disease

- Complete Response (CR): Complete disappearance of all target lesions with the exception of nodal disease. All target nodes must decrease to normal size (short axis < 10 mm). All target lesions must be assessed.

- Partial Response (PR): Greater than or equal to 30% decrease under baseline of the sum of diameters of all target measurable lesions. All target lesions must be assessed.

- Stable Disease (SD): Does not qualify for CR, PR or Progression. All target lesions must be assessed. Stable can follow PR only in the rare case that the sum increases by less than 20% from the nadir (smallest sum of diameters considering baseline and all assessments prior to the time point under evaluation), but enough that a previously documented 30% decrease no longer holds.

- Objective Progression (PD): 20% increase in the sum of diameters of target measurable lesions above the smallest sum observed (over baseline if no decrease in the sum is observed during therapy), with a minimum absolute increase of 5 mm.

- Not evaluable (NE): Progression has not been documented, and

    - one or more target measurable lesions have not been assessed; or

    - assessment methods used were inconsistent with those used at baseline; or

    - one or more target lesions cannot be measured accurately (eg poorly visible unless due to being too small to measure); or

    - one or more target lesions were excised or irradiated and have not reappeared or increased.

Non-target disease

- CR: Disappearance of all non-target lesions and normalization of tumor marker levels (if being followed). All lymph nodes must be 'normal' in size (<10 mm short axis).

- Non-CR/Non-PD: Persistence of any non-target lesions and/or tumor marker level (if being followed) above the normal limits.

- PD: Unequivocal progression of pre-existing lesions. Generally, the overall tumor burden must increase sufficiently to merit discontinuation of therapy. In the presence

of SD or PR in target disease, progression due to unequivocal increase in non-target disease should be rare.

- Not evaluable (NE): Progression has not been determined and
  - one or more non-target sites were not assessed; or
  - assessment methods were inconsistent with those used at baseline; or
  - one or more non-target lesions cannot be assessed (eg, poorly visible or unclear images); or
  - one or more non-target lesions were excised or irradiated and have not reappeared or increased.

New Lesions

The appearance of any new unequivocal malignant lesion indicates PD. If a new lesion is equivocal, for example due to its small size, continued assessment will clarify the etiology. If repeat assessments confirm the lesion, then progression should be recorded on the date of the initial assessment. A lesion identified in an area not previously scanned will be considered a new lesion.

Supplemental Investigations

- If CR determination depends on a residual lesion that decreased in size but did not disappear completely, it is recommended the residual lesion be investigated with biopsy or fine needle aspirate. If no disease is identified, objective status is CR.
- If progression determination depends on a lesion with an increase possibly due to necrosis, the lesion may be investigated with biopsy or fine needle aspirate to clarify status.

Subjective progression

Patients requiring discontinuation of treatment without objective evidence of disease progression should not be reported as PD on tumor assessment CRFs. This should be indicated on the end of treatment CRF as off treatment due to Global Deterioration of Health Status. Every effort should be made to document PD even after discontinuation of study treatment.

Non-Measurable Disease Only

The non-target lesion(s) assessment criteria apply when assessing non-measurable disease response. When assessing patients for an unequivocal progression, a useful test to apply is to consider if the increase in overall disease burden based on the change in non-measurable disease is comparable in magnitude to the increase that would be required to declare progressive disease for measurable disease: ie, an increase in tumor burden representing an

additional 73% increase in 'volume' (which is equivalent to a 20% increase diameter in a measurable lesion). In general, increase in tumor burden must be sufficient to require a change in therapy.

Objective Response Status at Each Assessment

Objective Response Status at Each Assessment for Patients With Measurable Disease at Baseline

Target Lesions	Non-Target Lesions	New Lesions	Objective Status
CR	CR	No	CR
CR	Non-CR/Non-PD or not all evaluated	No	PR
PR	Non-PD* or not all evaluated	No	PR
SD	Non-PD* or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes**	PD

*Non-PD includes CR and Non-CR/Non-PD
** New lesions must be unequivocal
CR = complete response; PD = progressive disease; PR = partial response; SD = stable disease; NE = not evaluable.

Objective Response Status at Each Assessment for Patients with Non-Measurable Disease at Baseline

Target Lesions	Non-target Lesions	New Lesions	Objective status
Not present	CR	No	CR
	Not all evaluated	No	NE
	Non-CR/Non-PD	No	Non-CR/Non-PD
	PD	Yes or No	PD
	Any	Yes**	PD

** New lesions must be unequivocal

Date of Response

If there are multiple scan dates associated with a tumor evaluation, ie, radiological assessments occur over a series of days rather than the same day, the earliest scan date associated with the evaluation will be used as the date of the assessment. The same convention should be used for the investigator assessment (IOTA CRF page).

Appendix 2. Modified RANO Criteria for Patients with CNS Malignancies

Adapted from *P Wen, et al: Response Assessment in Neuro-Oncology Clinical Trials. Journal of Clinical Oncology 35 (2017) 2439-49.*[15]

Methodology to Determine Tumor Measurement in CNS Tumors

Tumor dimensions are determined by measurement of the longest tumor dimension and its perpendicular for each target lesion. Regarding MRI imaging, the radiologist may select whatever sequence best highlights the tumor (T1 enhanced or T2 weighted or FLAIR images) and the same sequence should be used for serial measurements. Response determination will be based on a comparison of an area (W × T) between the baseline assessment and the study date designated in the follow-up Report Form. Reports for the follow-up exams should reiterate the measurements obtained at baseline for each target lesion. Non-target lesions or newly occurring lesions should also be enumerated in these reports, and changes in non-target lesions should be described.

Tumor response criteria are determined by changes in size using the longest tumor dimension, and its perpendicular. Either T1 or T2 weighted images are used - which ever gives the best estimate of tumor size. The following section describes the methodology. Longest diameter of target lesion(s) should be selected in the axial plane only for CT. For MRI imaging, the longest diameter can be measured from the axial plane or the plane in which the tumor is best seen or measured, provided the same plane is used in follow ups. This longest measurement of the tumor is referred to as the width (W).

The perpendicular measurements should be determined - transverse (T) measurement, perpendicular to the width in the selected plane. The cystic or necrotic components of a tumor are not considered in tumor measurements. Therefore, only the solid component of cystic/necrotic tumors should be measured. If cysts/necrosis compose the majority of the lesion, the lesion may not be “measurable”.

Options:

- If the cyst/necrosis is eccentric, the W and T of the solid portion should be measured, the cyst/necrosis excluded from measurement.
- If the cyst/necrosis is central but represents a small portion of the tumor (<25%), disregard and measure the whole lesion.
- If the cyst/necrosis is central but represents a large portion of the tumor, identify a solid aspect of the mass that can be reproducibly measured.

Leptomeningeal tumor spread is usually not a target lesion, and usually cannot be measured accurately. Presence and location of leptomeningeal tumor spread should be noted and change in extent/thickness assessed on follow up studies.

Selection of Target and Non-Target Lesions

For most CNS tumors, only one lesion/mass is present and therefore is considered a "target" for measurement/follow up to assess for tumor progression/response.

If multiple measurable lesions are present, up to 3 should be selected as "target" lesions. Target lesions should be selected on the basis of size and suitability for accurate repeated measurements.

All other lesions will be followed as non-target lesions (including CSF positive for tumor cells). The lower size limit of the target lesion(s) should be at least twice the thickness of the slices showing the tumor to decrease the partial volume effect (eg, 8 mm lesion for a 4 mm slice).

Any change in size of non-target lesions should be noted, though does not need to be measured.

Evaluation of Target Lesions (for T1-Gd and T2 Flair)

Response criteria are assessed in 2 dimensions – the product of W × T

To assess response/progression, the ratio is calculated:

W × T (current scan)
W × T (reference scan)

Development of new disease or progression in any established lesions is considered progressive disease, regardless of response in other lesions – eg, when multiple lesions show opposite responses, the progressive disease takes precedence.

- Complete Response (CR): Disappearance of all target lesions.

- Partial response (PR): ≥50% decrease in the sum of the products of the two perpendicular diameters of target lesions, compared to baseline measurement.

- Stable Disease (SD): Neither sufficient decrease in the sum of the products of the two perpendicular diameters of all target lesions to qualify for PR (taking as reference the initial baseline measurements), nor sufficient increase in a single target lesion to qualify for PD, (taking as reference the smallest disease measurement since the treatment started).

- Progressive Disease (PD): ≥25% increase in the sum of products of perpendicular diameters of ANY target lesion, taking as reference the smallest product observed since the start of treatment (see exception below); OR the appearance of one or more new lesions, OR worsening neurologic status not explained by causes unrelated to tumor progression (eg, anticonvulsant or corticosteroid toxicity, electrolyte

disturbances, sepsis, hyperglycemia, presumed post-therapy swelling etc.) PLUS any increase in tumor cross-sectional area (or tumor volume).

However, because radiotherapy may be associated with transient, reversible swelling during or within the first 3 months after completion of radiotherapy, there may be a lag time between the initiation of therapy and maximal antitumor effect. Removing a patient from protocol therapy as soon as the tumor area increases by 25%, may result in the treatment being terminated prematurely. It is quite possible that if these patients were maintained on protocol therapy, their disease might eventually stabilize and even regress.

Local progression is defined as progression of known residual tumor or the appearance of tumor at known prior sites of disease that were at some point without evidence of disease. Distant progression is defined as the appearance of tumor at sites other than known prior sites of disease. Distant progression most often occurs in the subarachnoid space and may occur at any point within the neuraxis. Although rare, extra-CNS metastasis represents distant failure. Combined local and distant progression is defined when evaluation of the entire neuraxis reveals local and distant progression.

Evaluation of Non-target Lesions (for T1-Gd and T2 Flair)

- Complete Response (CR): Disappearance of all non-target lesions.
- Incomplete Response/Stable Disease (IR/SD): The persistence of one or more non-target lesions.
- Progressive Disease (PD): The appearance of one or more new lesions and/or unequivocal progression of existing non-target lesions.

Overall Response Assessment

The overall response assessment takes into account response in both target and non-target lesion, the appearance of new lesions and neurologic exam, where applicable, according to the criteria described in the table below. The overall response assessment is shown in the last column, and depends on the assessments of target, non-target, new lesions, and neurologic exam in the preceding columns.

Overall Response Assessment for CNS Tumors (Modified RANO)				
Target lesions	**Non-target lesions**	**New lesions**	**Neurologic Exam**	**Overall response**
CR	CR	No	Improved, Clinically Significant OR Improved, Not Clinically Significant, Worsened, Not Clinically Significant, OR No Change OR Indeterminate	CR
CR	IR/SD	No	Improved, Clinically Significant OR Improved, Not Clinically Significant, Worsened, Not Clinically Significant, OR No Change OR Indeterminate	PR
PR	CR, IR/SD	No	Improved, Clinically Significant OR Improved, Not Clinically Significant, Worsened, Not Clinically Significant, OR No Change OR Indeterminate	PR
SD	CR, IR/SD	No	Improved, Clinically Significant OR Improved, Not Clinically Significant OR Worsened, Not Clinically Significant, OR No Change OR Indeterminate	SD
SD	CR, IR/SD	No	Worsened, Clinically Significant*	PD
PD	Any	Yes or No	Any	PD
Any	PD	Yes or No	Any	PD
Any	Any	Yes	Any	PD

Abbreviations: CR=complete response; IR=incomplete response; PD=progressive disease; PR=partial response SD=stable disease

* For patients with worsening neurologic exam not explained by causes unrelated to tumor progression, any evidence of progressive disease (does not require > 25% increase) constitutes an overall response of progressive disease

Appendix 3. Response Criteria for Neuroblastoma

Adapted from J.R. Park, et al: Revisions to the international neuroblastoma response criteria: A consensus statement from the National Cancer Institute clinical trials planning meeting. Journal of Clinical Oncology 35:2580-2587.[10]

This study will use the revised International Neuroblastoma Response Criteria for disease assessment. The updated response criteria incorporate current approaches to imaging of neuroblastoma, including functional imaging. Furthermore, a standardized approach to assessment of bone marrow involvement is included. The current INRC do not include methods of disease assessment that are less sensitive and/or specific for neuroblastoma (99Tc bone scan and catecholamine levels).

Key sites and terms

Primary site: The primary site will be identified as a measurable lesion ≥10 mm in diameter as assessed by cross sectional imaging (CT or MRI scan). Primary site measurements must be recorded in millimeters (or decimal fractions of centimeters). The longest diameter of the primary tumor will be recorded at baseline. Serial measurements of the primary tumor will include assessment of tumor size in the same orthogonal plane at the time of each evaluation. In patients with bilateral adrenal lesions, response will be based on the sum of the longest dimensions of both adrenal lesions unless biopsy proves one to be ganglioneuroma rather than neuroblastoma/ganglioneuroblastoma. In patients with multi-focal non-adrenal disease, the largest tumor will be considered the primary tumor. Response in additional lesions will be assessed as described below for metastatic lesions.

Tracer avidity (123I-MIBG or FDG-PET) in the primary site will be recorded at baseline. The scan appropriate for serial disease assessments should be used at each disease re-evaluation timepoint (eg, 123I-MIBG avid primary lesions should be followed using 123I-MIBG scans during therapy).

Malignant lymph nodes: To be considered pathologically enlarged and measurable, a metastatic lymph node must be ≥15 mm in short axis when assessed by CT or MRI scan (CT scan slice thickness recommended to be no greater than 5 mm). At baseline and in follow-up, only the short axis of a discreet lymph node will be measured and followed as per RECIST criteria. Patients with neuroblastoma may have conglomerate masses of non- discrete lymph nodes (ie, multiple contiguous retroperitoneal nodes). When a short axis of a discreet node cannot be identified, a lymph node conglomerate can be measured using the longest diameter of the composite lesion. Tracer avidity of metastatic nodes will be recorded at baseline and during disease evaluations.

For the purposes of response assessment, target lesions are disease sites that are measurable (non-nodal soft tissue mass ≥10 mm in longest dimension or lymph node ≥15 mm in short axis) and tracer avid OR are biopsy positive for neuroblastoma or ganglioneuroblastoma. The sum of diameters of target lesions is defined as the sum of the short axis of discrete lymph nodes (ie, cervical, axillary nodes) added to the sum of the longest diameters of non-lymph node soft tissue metastases.

Non-measurable disease: All other lesions (or sites of disease), including small lesions (longest diameter <10 mm or pathological lymph nodes with ≥10 to <15 mm short axis), are considered non-measurable disease. Bone lesions, leptomeningeal disease, ascites, pleural/pericardial effusions are considered non-measurable.

Bone lesions: Osteomedullary disease will be assessed using 123I-MIBG scans or FDG-PET scans. Technitium bone scans are no longer used as part of the revised INRC and are not included as part of disease reassessments during this trial. The extent of tracer avid disease will be evaluated using the Curie scoring system (see worksheet within this appendix). Single-photon emission computer tomography (SPECT) may be used to confirm the presence or absence of lesions in a given segment of the body. The absolute Curie score should be reported at baseline. A relative score (Curie score at the time of disease assessment divided by baseline Curie score) should be recorded at the time of each disease evaluation.

Bone marrow disease: Bilateral bone marrow aspirates and trephine biopsies are required at disease assessment timepoints. The extent of marrow involvement in all 4 samples should be recorded. Use of immunohistochemical staining for evaluation of trephine biopsies is strongly encouraged. The percentage of tumor infiltration of bone marrow space assessed by histologic evaluation of trephine/biopsies or counting the number of tumor cells in aspirates by cytology or immunocytology (recommended if available) divided by the number of hematopoietic /mononuclear cells evaluated to obtain a percentage involvement.

The bone marrow sample with the highest percentage of tumor infiltration is used for response assessment. If >0% to ≤ 5% tumor infiltration is the highest percentage seen among samples obtained, the result should be recorded as minimal marrow disease.

Response Criteria

PRIMARY (SOFT TISSUE) TUMOR RESPONSE[1]

RESPONSE	ANATOMICAL IMAGING + MIBG (FDG-PET2) IMAGING
Complete Response (CR)	• <10 mm residual soft tissue at primary site, AND • complete resolution of MIBG or FDG-PET uptake (for MIBG non-avid tumors) at primary site.
• **Partial Response (PR)**	• ≥30% decrease in longest diameter (LD) of primary site. • MIBG or FDG-PET uptake at primary site stable, improved or resolved.
• **Progressive Disease (PD)**	• >20% increase in longest diameter taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study), AND • a minimum absolute increase of 5 mm in longest dimension[3].
Stable Disease (SD)	Neither sufficient shrinkage for PR nor sufficient increase for PD at the primary site.

1. Not for use in assessment of metastatic sites.
2. For ^{123}I-MIBG non-avid tumors.
3. A mass that has not met PD measurement criteria but has fluctuating ^{123}I-MIBG avidity will not be considered progressive disease.

RESPONSE AT METASTATIC SOFT TISSUE AND BONE SITES

RESPONSE	ANATOMICAL IMAGING + MIBG (FDG-PET[1]) IMAGING

Complete Response (CR)	• Resolution of all sites of disease defined as:
	  • Non-primary target and non-target lesions measure <10 mm AND
	  • Lymph nodes identified as target lesions decrease to a short axis.
	  • <15 mm, AND
	  • MIBG uptake or FDG-PET uptake (for MIBG non-avid tumors) of non-primary lesions resolves completely.

Partial Response (PR)	• ≥30% decrease in sum of diameters[2] of non-primary target lesions compared to baseline, AND all of the following:
	  • Non-target lesions may be stable or smaller in size, AND
	  • No new lesions, AND
	  • ≥50% reduction in MIBG absolute bone score (Relative MIBG bone score ≥0.1 to ≤0.5) or ≥50% reduction in number of FDG-PET avid bone lesions.[3,4]

Progressive Disease (PD)	Any of the following:[5] • Any new soft tissue lesion detected by CT or MRI that is also MIBG avid or FDG-PET avid; • Any new soft tissue lesion seen on anatomic imaging that is biopsied and confirmed to be a neuroblastoma or ganglioneuroblastoma; • Any new bone site that is MIBG avid; • A new bone site that is FDG-PET avid (for MIBG non-avid tumors) AND has CT or MRI findings consistent with tumor OR has been confirmed histologically to be neuroblastoma or ganglioneuroblastoma; • >20% increase in longest diameter taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study), AND a minimum absolute increase of 5 mm in sum of diameters of target soft tissue lesions; • Relative MIBG score ≥1.2.[4]
Stable Disease (SD)	Neither sufficient shrinkage for PR nor sufficient increase for PD of non-primary lesions.

1. Used for MIBG non-avid tumors.
2. Sum of diameters is defined as the sum of the short axis of discrete lymph nodes (ie, cervical, axillary nodes) added to the sum of the longest diameters of non-lymph node soft tissue metastases. Masses of conglomerate non-discrete lymph nodes will be measured using longest diameter.
3. For patients with soft tissue metastatic disease, resolution of MIBG and/or FDG-PET uptake at the soft tissue sites is not required; all size reduction criteria must be fulfilled.
4. Relative Curie score is the absolute score for bone lesions at time of response assessment divided by the absolute score for bone lesions at entry onto a clinical trial. MIBG-SPECT or MIBG-SPECT/CT may be used for scoring purposes but the same imaging methodology should be used for all evaluations.
5. The post-infusion MIBG scan is not considered a diagnostic study for the purposes of response assessment. Progressive disease should NOT be designated on the basis of this scan.

BONE MARROW RESPONSE

RESPONSE	BONE MARROW STATUS[1]
Complete response (CR)	Bone marrow with no tumor infiltration upon reassessment, independent of baseline tumor involvement.
Progressive disease (PD)	Any of the following: • Bone marrow without tumor infiltration that becomes >5% tumor infiltration upon reassessment; or • Bone marrow with tumor infiltration that increases by >2 fold and has >20% tumor infiltration upon reassessment.
Minimal disease (MD)	Any of the following: • Bone marrow with ≤5% tumor infiltration and remains >0-≤5% tumor infiltration upon reassessment; or • Bone marrow with no tumor infiltration that becomes ≤5%. • tumor infiltration upon reassessment; or • Bone marrow with >20% tumor infiltration that has >0-≤5% tumor infiltration upon reassessment.
Stable disease (SD)	Bone marrow with tumor infiltration that remains positive with >5% tumor infiltration upon reassessment but does not meet CR, MD, or PD criteria.

1. Immunohistochemistry strongly encouraged.

DETERMINATION OF OVERALL RESPONSE

RESPONSE	CRITERIA
Complete Response (CR)	All components meet criteria for CR.
Partial Response (PR)	PR in at least 1 component and all other components are either CR, MD (Bone marrow), PR (Soft tissue or Bone) or Not involved (NI); no component with PD.
Minor Response (MR)	PR or CR in at least one component but at least one other component with SD; no component with PD.
Stable Disease (SD)	SD in one component with no better than SD or NI in any other component; no component with PD.
Progressive Disease (PD)	Any component with PD.

NI = Not involved, site not involved at study entry and remains not involved; MD = Minimal Disease, for bone marrow assessment only.

OVERALL BEST RESPONSE ASSESSMENT

Each patient will be classified according to his "best response" for the purposes of analysis of treatment effect.

Primary Tumor	Soft Tissue and Bone Metastatic Disease	Bone Marrow Metastatic Disease	Overall
CR	CR	CR	CR
CR for one response component with either CR or NI for other components			CR
CR	CR	MD	PR
CR	PR	CR	PR
CR	PR	MD	PR
CR	PR	NI	PR
CR	NI	MD	PR
PR	CR	CR	PR
PR	CR	NI	PR
PR	CR	MD	PR
PR	PR	CR	PR
PR	PR	NI	PR
PR	PR	MD	PR
PR	NI	CR	PR
PR	NI	NI	PR
PR	NI	MD	PR
NI	CR	MD	PR
NI	PR	CR	PR
NI	PR	MD	PR
CR	CR	SD	MR
CR	PR	SD	MR
CR	SD	CR	MR
CR	SD	MD	MR
CR	SD	SD	MR
CR	SD	NI	MR
CR	NI	SD	MR
PR	CR	SD	MR
PR	PR	SD	MR
PR	SD	CR	MR
PR	SD	MD	MR
PR	SD	SD	MR
PR	SD	NI	MR
PR	NI	SD	MR
SD	CR	CR	MR
SD	CR	MD	MR
SD	CR	SD	MR
SD	CR	NI	MR
SD	PR	CR	MR
SD	PR	MD	MR
SD	PR	SD	MR
SD	PR	NI	MR
SD	SD	CR	MR
SD	NI	CR	MR
NI	CR	SD	MR
NI	PR	SD	MR
NI	SD	CR	MR
SD	SD	MD	SD
NI	SD	MD	SD
SD	NI	MD	SD
NI	NI	MD	SD
SD	SD	SD	SD
SD	NI	SD	SD
SD	NI	SD	SD
NI	SD	SD	SD
NI	SD	NI	SD
NI	NI	SD	SD

CR= Complete Response; MD= Minimal Disease; PR: Partial Response; MR= Minor Response; SD= Stable Disease; PD= Progressive disease; NI= not involved; site not involved at study entry and remains not involved.

Appendix 4. Data Derivation Details

Enrollment/Randomization	Date of assignment of the enrollment (randomization) number
Study Day 1 (Phase 2 portion)	Randomization day
Study treatment start date	First date when a non-zero dose of study drug of respective treatment combination is administered
Day 1 (cycle start date; Phase 1 portion)	Day 1 of a cycle is every 21 days unless there is a dosing delay.
Cycle length (all but last cycle)	Cycle length is 21 days (previous cycle length may exceed planned length if there is a delay in study treatment administration).
Cycle X treatment delayed	If start of study treatment administration is delayed for Cycle X then Cycle X-1 is extended.
Last cycle	For patients off treatment, from Day 1 of final cycle to the last dose of study treatment. For patients on treatment, from Day 1 of most recent cycle start to protocol specified cycle length.
On-treatment period for TEAEs and other safety endpoints	From the date that the patient took the first dose of study treatment + 35 calendar days after the last dose of study treatment
Baseline value/assessment for safety and efficacy (Phase 1 portion)	Last available assessment prior to first dose of study treatment
Baseline value/assessment for efficacy (Phase 2 portion)	Last available assessment prior to randomization unless otherwise specified

Appendix 5. Study Treatment Modification and Compliance Derivation Details

Dose modifications

In the event of significant treatment-related toxicity, palbociclib dosing may be interrupted, reduced and/or delayed as described below. Dose reductions of IRN, TMZ or TOPO, CTX will be allowed depending on the type and severity of toxicity encountered.

Dose modifications may occur in 3 ways:

- Within a cycle: dosing interruption until adequate recovery and dose reduction, if required, during a given treatment cycle;
- Between cycles: next cycle administration may be delayed due to persisting toxicity when a new cycle is due to start;
- In the next cycle: dose reduction may be required in a subsequent cycle based on toxicity experienced in the previous cycle.

A **dose interruption** is defined as a planned dosing day with 0 mg administered within a cycle.

A **cycle delay** is defined as delay of the cycle start date based on the previous cycle start date, ie, if [start date of Cycle X – start date of Cycle (X-1) – 21] > 3.

A **dose reduction** is defined as a change to a non-zero dose level lower than the prior dose level (planned dose). Once a dose has been reduced for a given patient, all subsequent doses should be administered at that dose level, unless further dose reduction is required. The reduced doses in those subsequent cycles are not defined as dose reductions.

Summarizing Relative Total Dose (RTD) and Relative Total Dose Intensity (RTDI)

The dose of palbociclib/IRN/TMZ (mg) or palbociclib/TOPO/CTX (mg) will depend on the dose level (mg/m^2) that the patient is assigned to and will be adjusted for the patient's BSA (m^2) measured during a scheduled clinic visit at the start of a cycle.

The following types of summaries are proposed for administration of each study drug by treatment combination.

When palbociclib is administered orally (capsule or oral solution) QD on Days 1-14 of each 21-day cycle followed by 7 days off treatment (cyclical dosing) in combination with IRN (IV on Day 1-5 of each 21-day cycle) and TMZ (IV/orally on Day 1-5 of each 21-day cycle) and in combination with TOPO (IV on Day 1-5 of each 21-day cycle) and CTX (IV on Day 1-5 of each 21-day cycle), the following dose intensity summaries can be presented:

- RTDI for palbociclib, all cycles combined
- RTDI for IRN, all cycles combined
- RTDI for TMZ, all cycles combined
- RTDI for TOPO, all cycles combined

- RTDI for CTX, all cycles combined

The denominator for tables summarizing palbociclib, IRN or TMZ will be all patients who took at least a dose of palbociclib, IRN or TMZ, respectively; the denominator for tables summarizing palbociclib, TOPO or CTX will be all patients who took at least a dose of palbociclib, TOPO or CTX, respectively.

Conventions:

- Regular Cycle: There is a subsequent cycle after the current one.
- Last Cycle: The treatment is permanently discontinued after the current cycle.
- For Cycle X, actual cycle start date for each patient is
    - start date of initial dosing in the Cycle X Day 1, if the patient received study drug on that visit (ie, any study drug with dose>0 at that visit)
    - the first day of assessments in the Cycle X Day 1, if the patient did not receive any study drug on that visit (ie, dose=0 for any study drug at that visit). Use start date in the exposure page if available; if start date is not available then use date of collection of vital signs on Cycle X Day 1.
- For Cycle X, actual cycle end date for each patient is
    - for all cycles X except the last cycle, actual cycle end date = actual cycle (X+1) start date – 1 day;
    - for the last cycle, actual cycle end date = actual cycle start date + 21 – 1 day
- Intended Total Dose (mg) for a given cycle is calculated as Intended Daily Dose (mg; BSA adjusted) * Intended Dosed Days per cycle, where the intended dose level (mg/m^2) is fixed at the start of treatment rather than start of a cycle, and the last available BSA (m^2) of the patient during a scheduled clinic visit (ie, on or prior to the day of dosing) will be used (refer to Appendix 6, 7 and 8 for oral dosing for palbociclib and TMZ, respectively).
- Intended Dosed Days Per Cycle for palbociclib
    - 14 days for a regular cycle, or
    - Minimum of (14 days, actual treatment duration) for the last cycle
- Intended Dosed Days Per Cycle for IRN, TMZ, TOPO or CTX
    - 5 days for a regular cycle, or
    - Minimum of (5 days, actual treatment duration) for the last cycle

- Intended Treatment Duration is the same for the entire dosing period, except for the last cycle which is from the first dose date of the last cycle to the last date of the last cycle or duration of 21 days, whichever is shorter, ie,
    - 21 days for a regular cycle, or
    - Minimum of (21 days, actual treatment duration) for the last cycle
- Actual Total Dose (Cumulative Dose) (all cycles combined) is the total dose a patient actually took over the treatment duration of n cycles and accounts for dose interruption, dose reduction and overdose.
- Actual Treatment Duration (all cycles combined) is the actual treatment duration over the n cycle including cycle delays.

Cyclical Palbociclib and IRN/TMZ/TOPO/CTX, All Cycles Combined

Study Drug	Calculation of RTD/RTDI
Cyclical Palbociclib (2/1 QD in a 21-day cycle)	$RTD = \frac{Actual\ Total\ Dose}{Intended\ Total\ Dose} * 100\%$ *Actual Total Dose (mg)* = Sum over all cycles of the actual total dose taken over a cycle *Intended Total Dose (mg)* = (Intended Total Dose per cycle) * (Total number of cycles per CRF) Note: Calculation of RTD is optional.
	$RTDI = \frac{Actual\ Total\ Dose\ Intensity}{Intended\ Total\ Dose\ Intensity} * 100\%$ *Actual Total Dose Intensity (mg/day)* = (Sum of overall Actual Total Dose) / (Sum of overall Actual Treatment Duration) *Intended Total Dose Intensity (mg/day)* = (Sum of overall Intended Total Dose during the study) /(Sum of overall Intended Treatment Duration)
Cyclical IRN/TMZ/TOPO/ CTX (Day 1- 5 QD in a 21-day cycle)	$RTD = \frac{Actual\ Total\ Dose}{Intended\ Total\ Dose} * 100\%$ *Actual Total Dose (mg)* = Sum over all cycles of the actual total dose taken over a cycle

Study Drug	Calculation of RTD/RTDI
	Intended Total Dose (mg) = (Intended Total Dose per cycle) * (Total number of cycles per CRF) Note: Calculation of RTD is optional
	$RTDI = \frac{Actual\ Total\ Dose\ Intensity}{Intended\ Total\ Dose\ Intensity} * 100\%$ *Actual Total Dose Intensity* = (Sum of overall Actual Total Dose) / (Sum of overall Actual Treatment Duration) *Intended Total Dose Intensity* = (Sum of overall Intended Total Dose during the study) /(Sum of overall Intended Treatment Duration)

Appendix 6. Palbociclib Oral Capsule Dosing Table

Body Surface Area (m^2)*	Actual Daily Dose (mg)				
	40 mg/m^2	**55 mg/m^2**	**75 mg/m^2**	95 mg/m^2	115 mg/m^2
0.3	Oral Solution needed**	Oral Solution needed	Oral Solution needed	Oral Solution needed	Oral Solution needed
0.4	Oral Solution needed	Oral Solution needed	Oral Solution needed	Oral Solution needed	Oral Solution needed
0.5	Oral Solution needed	Oral Solution needed	Oral Solution needed	Oral Solution needed	Oral Solution needed
0.6	Oral Solution needed	Oral Solution needed	Oral Solution needed	Oral Solution needed	75
0.7	Oral Solution needed	Oral Solution needed	Oral Solution needed	Oral Solution needed	75
0.8	Oral Solution needed	Oral Solution needed	Oral Solution needed	75	100
0.9	Oral Solution needed	Oral Solution needed	75	Oral Solution needed***	100
1.0	Oral Solution needed	Oral Solution needed	75	100	125
1.1	Oral Solution needed	Oral Solution needed	75	100	125
1.2	Oral Solution needed	75	100	125	150
1.3	Oral Solution needed	75	100	125	150
1.4	Oral Solution needed	75	100	125	150
1.5	Oral Solution needed	75	100	125	175
1.6	75	100	125	150	200
1.7	75	100	125	150	200
1.8	75	100	125	175	200
1.9	75	100	125	175	200
2.0	75	100	150	200	200
2.1	75	125	150	200	200
2.2	100	125	150	200	200
2.3	100	125	150	200	200
2.4	100	125	200	200	200

Note: Highest dose is capped at 200 mg, the palbociclib monotherapy MTD in adults on the Schedule 2/1 (dosing for 14 days, followed by 7 days off treatment).

*The BSA should be rounded to nearest tenth m^2 for dose determination.

**Oral solution must be used when a patient's BSA-based dose is more than 15% below the 75-mg capsule strength.

*** Oral solution needed due to overlap in capsule strength.

Appendix 7. Volume of Palbociclib Oral Solution to be Administered by Dose and BSA

Palbociclib Oral Solution Administration Table

Body Surface Area (m^2)	Dose Level - 40 mg/m^2		Dose Level – 55 mg/m^2		Dose Level – 75 mg/m^2		Dose Level – 95 mg/m^2		Dose Level – 115 mg/m^2	
	Daily Dose Delivered (mg)	Oral Solution Volume Delivered (mL)	Daily Dose Delivered (mg)	Oral Solution Volume Delivered (mL)	Daily Dose Delivered (mg)	Oral Solution Volume Delivered (mL)	Daily Dose Delivered (mg)	Oral Solution Volume Delivered (mL)	Daily Dose Delivered (mg)	Oral Solution Volume Delivered (mL)
0.3	12.5	0.5	17.5	0.7	22.5	0.9	30	1.2	35	1.4
0.4	15	0.6	22.5	0.9	30	1.2	40	1.6	45	1.8
0.5	20	0.8	30	1.2	40	1.6	50	2	60	2.4
0.6	25	1	35	1.4	45	1.8	55	2.2	70	2.8
0.7	25	1	40	1.6	55	2.2	65	2.6	80	3.2
0.8	30	1.2	45	1.8	60	2.4	75	3	90	3.6
0.9	35	1.4	50	2	70	2.8	85	3.4	105	4.2
1	40	1.6	55	2.2	75	3	95	3.8	115	4.6
1.1	45	1.8	60	2.4	85	3.4	105	4.2	125	5
1.2	50	2	65	2.6	90	3.6	115	4.6	140	5.6
1.3	50	2	70	2.8	100	4	125	5	150	6
1.4	55	2.2	75	3	105	4.2	135	5.4	160	6.4
1.5	60	2.4	85	3.4	115	4.6	145	5.8	175	7
1.6	65	2.6	90	3.6	120	4.8	150	6	185	7.4
1.7	70	2.8	95	3.8	130	5.2	160	6.4	195	7.8
1.8	70	2.8	100	4	135	5.4	170	6.8	200	8
1.9	75	3	105	4.2	145	5.8	180	7.2	200	8
2	80	3.2	110	4.4	150	6	190	7.6	200	8
2.1	85	3.4	115	4.6	160	6.4	200	8	200	8
2.2	90	3.6	120	4.8	165	6.6	200	8	200	8
2.3	90	3.6	125	5	175	7	200	8	200	8
2.4	95	3.8	130	5.2	180	7.2	200	8	200	8

Note: Highest dose is capped at 200 mg, the palbociclib monotherapy MTD in adults on the 2/1 Schedule (dosing for 2 weeks, followed by 1 week off treatment). Palbociclib oral solution will be administered using 1 mL or 5 mL oral syringes. Volumes >1 mL will be administered using a 5 mL syringe (marked increments of 0.2 mL), while volumes ≤1 mL will be administered using a 1 mL oral syringe (marked increments of 0.1 mL).

Appendix 8. Dosing Nomogram for Oral Temozolomide

Body Surface Area (m^2)	Actual Daily Dose (mg)	
	100 mg/m^{2*}	75 mg/m^{2**}
0.3	30	20
0.4	40	30
0.5	50	35
0.6	60	45
0.7	70	50
0.8	80	60
0.9	90	65
1.0	100	75
1.1	110	80
1.2	120	90
1.3	130	95
1.4	140	105
1.5	150	110
1.6	160	120
1.7	170	125
1.8	180	135
1.9	190	140
2.0	200	150
2.1	210	155
2.2	220	165
2.3	230	170
2.4	240	180

Use combination of capsules (5 mg, 20 mg, 100 mg) for actual daily dose.
* Starting dose
** Dose reduction per protocol

Appendix 9. Palatability Questionnaire

Palbociclib Oral Solution Taste Feedback Questionnaire

Questionnaire should be administered to children or adults, preferably by the clinician or the nurse at the time of the first dose. After that, the questionnaire should be completed on Day 14 of Cycle 1, and Day 1 and 14 of Cycle 2.

Questions need to be answered shortly after dose administration (ideally within 10 minutes) so the data is collected within attention span of young children.

Please gather the following background information:

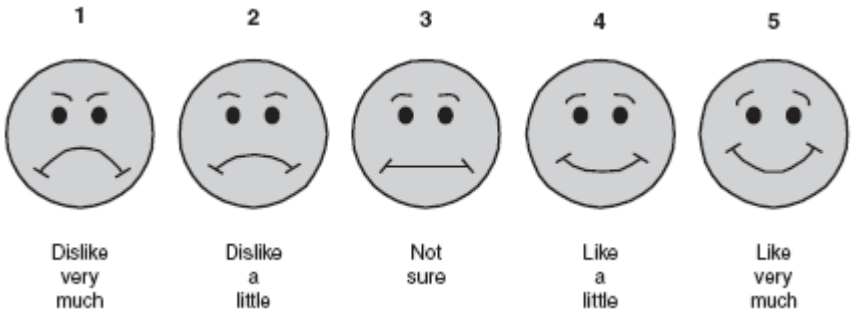
Background Information

Study/Study Site	
Patient ID	
Patient age	
Treatment Cycle Number and Day of Cycle	
Date/Time of administration of medicine	
Questionnaire completed by (circle one)	Patient Nurse Caregiver such as parent or guardian
What volume of oral solution is taken daily?	

Questionnaire

1. Did you drink (or eat) water (or something else) after you took the medicine before answering these questions? If you did what was it?

    1a) How much do you like this medicine overall?



2. Was dosing successful (entire dose administered and swallowed)? Please circle one (YES or NO)

   2a) If the answer to question 2 is “YES” , then based on the overall taste - did you have trouble swallowing the medicine?





Yes, a lot of trouble **Not sure** **No, very easy to swallow**

   2b) If the answer to question 2 is “NO” please indicate the reason for unsuccessful dosing/not being able to swallow all of the medicine?

Appendix 10. List of Abbreviations

Abbreviation	Term
AE	adverse event
ASCO	American Society of Clinical Oncology
AUC_{τ}	area under the concentration-time curve at steady state over the dosing interval τ
$AUC_{ss,\tau}$	area under the concentration-time curve during a dosing interval τ at steady state
BIRC	blinded independent review committee
BLQ	below the level of quantification
BMI	body mass index
BOR	best overall response
BSA	body surface area
CI	confidence interval
C_{max}	maximum plasma concentration
CL/F	oral plasma clearance
CNA	circulating nucleic acid
CNS	central nervous system
CR	complete response
CRF	case report form
eCRF	electronic case report form
CSR	clinical study report
$C_{ss,max}$	maximum observed plasma concentration during a dosing interval at steady state
$C_{ss,trough}$	measured concentration at the end of a dosing interval at steady state (taken directly before next administration)
CTCAE	Common Terminology Criteria for Adverse Events
CTX	cyclophosphamide
CV	coefficient of variation
DL	dose level
DLT	dose limiting toxicity
DNA	deoxyribonucleic acid
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EFS	event-free survival
EWS	Ewing sarcoma
FAS	full analysis set
FSH	follicle stimulating hormone
G-CSF	ganulocyte-colony stimulating factor
HbA1c	hemoglobin A1c
HR	hazard ratio *or* heart rate
INR	international normalized ratio
INRC	International Neuroblastoma Response Criteria
IP	investigational product

Abbreviation	Term
IRN	irinotecan
IV	intravenous
LH	luteinizing hormone
MAD	maximum administered dose
MD	minimal disease *or* multiple dose
MedDRA	Medical Dictionary for Regulatory Activities
MIBG	metaiodobenzylguanidine
MR	minor response
MTD	maximum tolerable dose
NE	not evaluable
NA	not applicable
NC	not calculated
NCI	National Cancer Institute
OR	objective response
ORR	objective response rate
OS	overall survival
PD	pharmacodynamic(s) *or* progressive disease
PFS	progression-free survival
PK	pharmacokinetic(s)
PR	partial response
PT	prothrombin time
QD	once daily
QT	time between the start of the Q wave and the end of the T wave
QTcF	QT interval corrected using Fridericia's formula
QTcB	QT interval corrected using Bazett's formula
RANO	Response Assessment in Neuro-Oncology
RB	retinoblastoma protein
RECIST	Response Evaluation Criteria in Solid Tumours
ROW	Rest of the world
RP2D	recommended Phase 2 dose
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease *or* standard deviation
TEAE	treatment-emergent adverse event
TK	thymidine kinase
T_{max}	maximum plasma concentration time
TMZ	temozolomide
TOPO	topotecan