Statistical Analysis Plan

Protocol Title:	**Phase 1b/2 Study of Carfilzomib in Combination With Induction Chemotherapy in Children With Relapsed or Refractory Acute Lymphoblastic Leukemia**
Short Protocol Title:	A Phase 1b/2 study of Carfilzomib in Pediatric ALL
Protocol Number:	20140106 (formerly CFZ008)
NCT Number:	NCT02303821
Authors:	[redacted], **Ph.D., MS** **Biostatistics Senior Manager, Amgen Inc** [redacted], **Ph.D.** **Biostatistician, PAREXEL International** [redacted], Ph.D. Biostatistics Director, Amgen Inc
Sponsor:	Onyx Therapeutics, Inc., an Amgen Inc. subsidiary One Amgen Center Drive Thousand Oaks, California 91320 United States Phone: +1-805-447-1000 Fax: +1-805-480-4978
SAP Date:	**Document Version** — **Date** Original (v [1.0]) — 09 January 2019 **Amendment 1 (v [2.0])** — **21 August 2020** Amendment 1 (v [2.0]) — 15 March 2021 **Amendment 2 (v [3.0])** — **29 August 2023**

Version Number	Date (DDMMMYYY)	Summary of Changes, including rationale for changes
Original (v1.0)	09 January 2019	Original Version
Amendment 1 (v2.0)	21 August 2020	Rationale: The purpose of this SAP amendment is for: The protocol amendment of adding a phase 2. The phase 2 externally-controlled study is a multicenter, single-group study of carfilzomib in combination with VXLD in pediatric subjects. Changes: List of Abbreviations and Definitions of Terms 1.Introduction Added: phase 1b and 2 protocol for study 20140106, AMG 981/Carfilzomib dated 07August2020. This SAP also provides details of utilizing an external control arm from the observational study 20180065 to be used for the analysis of phase 2 in study 20140106 after propensity score adjustment. The details are included in Appendix B. The scope of this document includes the interim analysis, the primary analysis and the final analysis that are planned and will be executed by the Amgen Global Biostatistical Science department unless otherwise specified, e.g., standard PK analysis may be provided by Clinical Pharmacology, Modeling and Simulation (CPMS). Appendix B: Added: Phase 2 Statistical Analysis Plan
	19 Nov 2020	Per FDA's feedback Updated: Table 23.1, Number of prior lines of salvage therapy is treated as categorical variable [redacted] Added: Handling of missing data for covariates in Section 20.2

Version Number	Date (DDMMMYYY)	Summary of Changes, including rationale for changes
		Section 15.2 additional detail of sample size calculation [redacted] Supportive analysis for primary endpoint of CR to descriptively compare with the best available therapy via literature search in Section 21.5.1
	15 Mar 2021	Per protocol amendment V9.0 Removed: In Section 21.5, Statistical comparison between experimental arm and external control arm of duration of response was removed and only descriptive statistics will be summarized. Updated: Details about the role of the independent biostatisticians who will support the futility efficacy analysis for Data Monitoring Committee (DMC) in Section **21.1**. Handling of missing data for adverse events, concomitant medications, death, and anti-cancer therapy are replaced by Amgen standard imputation rules in Appendix F. Per FDA request, the number of pre-specified covariates for the primary analysis was reduced. Some definitions were updated for administrative reasons. Added: Added 3 additional subgroups per protocol amendment V9.0: chemotherapy versus immunotherapy as last therapy prior to enrollment, anthracycline omitted from VXLD in induction (yes vs no), geographical region (EU vs Not EU) in Section 16.2.

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
Amendment 2 (v3.0)	**29 Aug 2023**	Model selection algorithm has been removed from the primary analysis and added as a sensitivity analysis. The planned analyses for echocardiograms were added in Section 21.6.8. Added COVID-19 related analysis in Section 21.2, 21.3 and 21.6.10 **Major changes to align with protocol amendment 10/11 and 12:** **1. Section 16. Objectives and Endpoints:** • **added "treatment response will be derived by the sponsor based upon local evaluation of bone marrow, peripheral blood, and extramedullary disease status."** • **reworded primary endpoint as "CR after induction therapy"** • **reworded the secondary endpoint of response after consolidation therapy to "CR, CRp, CRh, and CRi after consolidation therapy"** • [redacted] **2. Section 17.2. Sample Size:** • **changed sample size for CFZ-VXLD from "approximately 40 B-cell and 60 T-cell ALL subjects" to "a minimum of 50 B-cell ALL subjects and a minimum of 30 T-cell ALL subjects"** • **changed sample size for external control from "150 B-cell ALL subjects and 70 T-cell ALL subjects" to "approximately 74 B-cell ALL subjects and approximately 60 T-cell ALL" subjects"** • **updated the numbers in the table for sample size and power estimates accordingly** • **changed the number of subjects in experimental arm at interim analysis (IA) from 50 to 60** **3. Section 18.2. Subgroup:** • **added subgroup "History of primary induction failure"**

Version Number	Date (DDMMMYYY)	Summary of Changes, including rationale for changes
		• changed the category of "Relapse time from prior allogeneic HSCT" to <12 months, ≥12 months, Unknown • clarified age subgroup to Age at diagnosis and changed the categories to <1 year, 1 to <2 years, 2 to <10 years, ≥10 years • added subgroup Sex, Isolated bone marrow at relapse, Blast percent in bone marrow prior to the qualifying treatment • changed geographical region category **4. Section 19. Definitions:** • added a table to list all the study treatments • changed the calculation of Actual Dose Intensity and Relative Dose Intensity by using treatment days instead of duration of treatment (days) • changed the term 'Disease response' to 'Treatment response', and changed 'complete response' to 'complete remission'; also changed the central lab for response evaluation to sponsor derived response per local evaluation • updated the description of definition for EFS • added the definition for CR rate, rate of CRi or better, last known alive date, history of primary induction failure • updated the Study Day 1 definition to keep consistent with protocol • added definition for MRD endpoints • updated the definition of treatment response **5. Section 20. Analysis Sets:** • removed "Full Analysis Set" • added Induction Safety Analysis Set and Consolidation Safety Analysis Set • added the IPDs that will be excluded from Per Protocol Set • updated the Interim Futility Analysis Set

Version Number	Date (DDMMMYYY)	Summary of Changes, including rationale for changes
		• **removed Earliest Relapse Analysis Set** **6. Section 21. Planned Analyses:** • **reorganized the paragraphs and updated the language for Interim Analysis section** • **changed the sample size required in experimental arm at interim analysis from "approximately first 50% (20 with B-cell and 30 with T-cell)" to "at least 60 subjects"** • **changed that the interim analysis will be performed by the IBG** • **added "at least 100 enrolled" subjects in the experimental arm "have received at least one dose of carfilzomib" to Primary Analysis** **7. Section 22.3. Handling of Missing and Incomplete Data** • **added initial ALL diagnosis, prior ALL therapy medication, prior HSCT, remission to prior ALL therapy, and relapse to prior ALL therapy** **8. Section 23.3. Subject Accountability** • **added MRD sample disposition** **9. Section 23.5. Demographic and Baseline Characteristics** • **added additional age groups** • **added an additional demographics summary presenting side-by-side for phase 1 total, phase 2 total, and total of phase 1 and phase 2 to better characterize the demographics** **10. Section 23.6. Efficacy Analysis:** **updates in the efficacy endpoint summary tables and in the text:** **(a) specified that the descriptive summary of CR and CRi or better at the end of Induction, EFS, OS, and DOR will also be performed based on Safety Analysis Set and Per Protocol Set** **(b) removed sIPTW from the sensitivity analysis since the treatment effect is estimated based on the average treatment effect for treated (ATT)**

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
		subjects per protocol, whereas sIPTW is based on the average treatment effect (ATE) for the entire population **(c) removed the PS matching from the sensitivity analysis due to small sample size of B-cell and T-cell in each arm and an expected smaller sample size after PS matching** **(d) removed the Earliest Relapse Analysis Set from the sensitivity analysis due to small sample size for external control** **(e) added investigator assessed treatment response, and sponsor derived treatment response using central lab data for the sensitivity analysis** [redacted] **11. Section 23.7. Safety Analysis:** • **specified TEAEs definition for induction treatment period and consolidation treatment period** • **added EOI (event of interest) and treatment related SAEs to the analysis** • **removed lab analyte LDH and the table for the list of all lab tests** • **added hydrocortisone to the list of CNS combination drugs** • **specified (1) exposure summary will be separated for induction treatment period and consolidation treatment period; (2) exposure summary for IT chemotherapy, only includes number of doses administered, and ratio of number of actual doses to number of the protocol specified doses; (3) the number and percentage of subjects receiving VXLD or an approved alternate agent in induction chemotherapy will be summarized for each agent; (4) specified that the summary of total dose received, average dose per administration, actual dose intensity and relative dose**

Version Number	Date (DDMMMYYY)	Summary of Changes, including rationale for changes
		intensity will not include the alternative agents of VXLD **12. Section 23.8. Other Analyses** [redacted] **15. Appendix F. Handling of Incomplete Dates and Missing Dates** • **added F4. Imputation for Date of Initial ALL Diagnosis** • **added F5. Imputation for Start Date and End Date of Prior ALL Therapy Medication, Date of Prior Hematopoietic Stem Cell Transplant, Date of Remission (CRi or Better) to Prior ALL Therapy,**

Version Number	Date (DDMMMYYY)	Summary of Changes, including rationale for changes
		and Date of Relapse to Prior ALL Therapy **16. Appendix G. Treatment Response Definitions** • **added this section to specify the definitions/derivations for treatment response of CR, CRp, CRh, CRi, and for PD, SD and NE** **17. The terminology complete response was changed to complete remission throughout the SAP.**

1. Table of Contents

List of Table

List of Abbreviations and Definition of Terms

Abbreviation or Term	Definition/Explanation
AE	Adverse Event
ANC	Absolute Neutrophil Count
ALL	Acute Lymphoblastic Leukemia
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
AUC	Area Under Curve
BLR	Bayesian Logistic Regression
BFM	Berlin-Frankfurt--Münster
BSA	Body Surface Area
BUN	Blood Urea Nitrogen
CAR-T cell	Chimeric Antigen Receptor T cell
CAS	**Chemotherapy Analysis Set**
CBC	Complete Blood Count
CFZ	Carfilzomib
CI	Confidence Interval
Cmax	Maximum Plasma Concentration
CNS	Central Nervous System
COG	Children's Oncology Group
CR	Complete **Remission**
CRh	Complete **Remission** with partial hematological recovery
CRi	Complete **Remission** with incomplete hematologic recovery
CRp	Complete **Remission** without Platelet Recovery
CRR	Complete **Remission** Rate
CSRC	Cohort Safety Review Committee
DE 1	Dose Escalation Cohort 1
DE 2	Dose Escalation Cohort 2
DLT	Dose Limiting Toxicity
DMC	Data Monitoring Committee
DMP	Data Management Plan
DOR	Duration of Response
ECG	Electrocardiogram
ECHO	**Echocardiogram**
EFS	Event Free Survival

eCRF	Electronic Case Report Form
GSO-DM	Global Study Operations – Data Management
IA	Interim Analysis
IBG	Independent Biostatistics Group
ID	Induction Death
IPTW	Inverse Probability Treatment Weights
IT	Intrathecal
MedDRA	Medical Dictionary for Regulatory Activities
MRD	Minimal Residual Disease
MTD	Maximum Tolerated Dose
NCI-CTCAE	National Cancer Institute - Common Terminology Criteria for Adverse Events
NE	Not Evaluable
NGS	Next Generation Sequencing
NCRM	New Continual Reassessment Method
OR	Odds ratio
ORR	Overall Response Rate
OS	Overall Survival
PAS	**Primary Analysis Set**
PBMC	Peripheral Blood Mononuclear Cells
PCR	Polymerase Chain Reaction
PD	Progressive Disease
[redacted]	
PK	Pharmacokinetics
PoCR	Probability of CR
PR	Partial Response
PS	Propensity Score
R3 backbone	Dexamethasone, Mitoxantrone, PEG asparaginase, Vincristine
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SCT	Stem Cell Transplant
SD	Stable Disease; Standard Deviation
SoC	Standard of Care
TEAE	Treatment-emergent Adverse Event
Tmax	Time to maximum plasma concentration
TMP-SMX	Trimethoprim-Sulfamethoxazole

VAS	VXLD/VPLD Analysis Set
VIF	Variance Inflation Factor
VXLD	Daunorubicin, Dexamethasone, PEG asparaginase, Vincristine

2. Introduction

The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within the phase 1b and 2 protocol for study 20140106, AMG 981/Carfilzomib dated **28 August 2023**. This SAP also provides details of utilizing an external control arm from the observational study 20180065 to be used for the analysis of phase 2 in study 20140106 after propensity score adjustment. The details are included in Appendix B. The scope of this document includes the interim analysis, the primary analysis and the final analysis that are planned and will be executed by the Amgen Global Biostatistical Science department unless otherwise specified, e.g., standard PK analysis may be provided by Clinical Pharmacology, Modeling and Simulation (CPMS).

3. Phase 1 Objectives, Endpoints and Hypotheses

4. Objectives and Endpoints

Objectives	Endpoints
Primary	
To assess the safety and tolerability of carfilzomib, alone and in combination with induction chemotherapy, for the treatment of children with relapsed or refractory ALL	Safety and tolerability of carfilzomib alone and in combination with induction chemotherapy as defined by the type, incidence, severity, and outcome of adverse events (AEs); changes from baseline in key laboratory analytes, vital signs, and physical findings. Time to toxicity will be evaluated to differentiate single-agent carfilzomib from carfilzomib in combination with induction chemotherapy
To determine the MTD of carfilzomib in combination with induction chemotherapy and to recommend a phase 2 dose of carfilzomib in combination with induction chemotherapy	Incidence of DLT (definition of DLT per protocol Section 8.3)
Secondary	
To evaluate the combined rate of bone marrow CR and bone marrow CRp at the end of the Induction Cycle	Combined proportion of subjects who achieve CR or CRp at the end of the Induction Cycle
To characterize the PK of carfilzomib alone and in combination with induction chemotherapy	Pharmacokinetic parameters, principally maximum plasma concentration (C_{max}) and area under the curve (AUC), alone and in combination with induction chemotherapy, derived from levels of carfilzomib assayed in PK samples
To estimate the proportion of subjects who achieve MRD status $< 10^{-3}$ and $< 10^{-4}$ lymphoblasts at the end of the Induction Cycle	Proportion of subjects who achieve Minimal Residual Disease (MRD) status $< 10^{-3}$ and $< 10^{-4}$ lymphoblasts at the end of the Induction

	Cycle as assessed by next generation sequencing (NGS)

Exploratory
[redacted]

4.1 Hypotheses and/or Estimations

The number of DLTs and AEs at each dose level will be summarized separately for the carfilzomib single-agent Lead-in Window and the Induction Cycle. Adverse events will also be summarized separately for subjects who receive the optional cycle of consolidation chemotherapy.

5. Phase 1 Study Overview

5.1 Study Design

This is a nonrandomized, multicenter, Phase 1b dose-escalation and dose-expansion study of carfilzomib in combination with induction chemotherapy, comprising either an R3 backbone of dexamethasone, mitoxantrone, PEG-asparaginase, and vincristine as Dose Escalation 1 (DE 1) or a VXLD backbone of vincristine, dexamethasone,

PEG-asparaginase and daunorubicin as Dose Escalation 2 (DE 2) in children with relapsed or refractory ALL. During the DE 1 portion of the study only, the Induction Cycle will be preceded by a 1-week carfilzomib single-agent Lead-in Window. Subjects in both dose escalation portions of the study will receive a 4-week cycle of induction chemotherapy and have the option to receive a 4-week cycle of consolidation chemotherapy, if stable disease or better response is achieved at the end of the Induction Cycle.

The dose-escalation design will use the Bayesian 2-parameter logistic regression model, the new continual reassessment method (NCRM) applied to observed dose limiting toxicities (DLTs) occurring during the Lead-in Window and Induction Cycle (Neuenschwander 2008). The definition of DLT is provided in Section 8.3 of the protocol.

Before starting a new cohort, the data available from the DLT evaluable subjects are applied to an algorithm that computes the posterior distribution of the DLT rate at each dose level using the prespecified prior distributions. The evaluation of posterior probabilities of the estimated DLT rate at each dose will be used in the dose-escalation decision and the final determination of the MTD: the target toxicity interval is (20% – 33%) and the excessive/unacceptable toxicity interval is (> 33% – 100%).

A cohort size of 2 will be used in DE 1. The starting dose is 20 mg/m^2; the maximum planned dose and the minimum planned dose are 45 mg/m^2 and 20 mg/m^2, respectively. One or more cohorts of 2 subjects may be enrolled to each dose level, depending on the observed toxicity. After every cohort has completed the 4-week Induction Cycle, the Cohort Safety Review Committee (CSRC) will review the safety data and dose level recommended by the algorithm. The CSRC will decide on expanding the dose level by 2 subjects, advancing to the next dose level, de-escalating to a lower dose level, or determining the MTD. Unlike for de-escalation decisions, no dose level skipping will be allowed in dose escalation decisions. The algorithm will stop when the sample size reaches 18 or when the DLT rate of the recommended MTD has a 95% posterior credible interval within the pre-specified range of 5% to 60%.

A cohort size of 3 will be used in DE 2. The starting dose level is 27 mg/m^2, while the planned dose levels are 27, 36, 45, 56 and 70 mg/m^2. One or more cohorts of 3 subjects may be enrolled to each dose level, depending on the observed toxicity. After every cohort has completed the 4-week Induction Cycle, the CSRC will review the safety data and dose level recommended by the algorithm. The CSRC will decide on expanding the

dose level by 3 subjects, advancing to the next dose level, de-escalating to a lower dose level or determining the MTD. Unlike for de-escalation decisions, no dose-level skipping will be allowed in dose-escalation decisions.

The algorithm will stop when the sample size reaches 24 or if both of the first two criteria, or the third criterion are met:

- The probability of the Target Toxicity interval at the candidate MTD exceeds 40%

  AND

- A minimum of 2 cohorts (6 subjects) are accrued and treated at the candidate

  MTD

  OR

- The algorithm recommends a dose level that has been recommended twice and with 6 subjects treated, and the CSRC agrees with the recommendation

If clinically justified, CSRC may halt enrollment and recommend the MTD or the dose for the phase 2 study based on available data and the rules of the CSRC charter (See protocol Section 13.4).

A minimum of 2 subjects (1 cohort) will be enrolled at the starting dose level in DE 1. A minimum of 3 subjects (1 cohort) will be enrolled at the starting dose level in DE 2.

After each update from the NCRM algorithm, the results will be evaluated for the recommendation. The next dose level will be the dose with the highest posterior probability of having a toxicity rate in the target toxicity interval, (i.e., 20% – 33%), and under the constraint that the probability of excessive/unacceptable toxicity (i.e., $>$ 33% – 100%), is less than 40%.

The MTD will be selected based on the above criteria or the stopping rules as specified in Section 5.2 of protocol and the recommendation of the CSRC. When the protocol conditions for ending the study are met, the CSRC will review the data and recommend on a dose for the Phase 2 study.

5.2 Sample Size

The estimated sample size is based on the NCRM algorithm. A minimum 1 cohort (of 2 DLT-evaluable subjects) will be enrolled before the trial stops in the DE 1 portion of the study. A minimum 1 cohort (of 3 DLT-evaluable subjects) will be enrolled before the trial stops in the DE 2 portion. Based on previous simulation results, the algorithm generally finds the MTD with reasonable probability under various dose toxicity

distribution and prior assumptions with a maximum of 18 subjects in each dose escalation.

Initially, approximately 18 subjects were planned to participate in each of the 2 dose escalations. After revising the protocol to add 2 dose levels, the sample size for VXLD portion has been expanded to 24 subjects. Additional subjects may be enrolled to treat up to 12 subjects at the MTD. A minimum of 3 subjects are targeted to be enrolled in each of the following age groups at the time of their enrollment onto this study:

- 12 to < 24 months old
- 2 to < 12 years old
- 12 to 18 years old

A minimum number for subject enrollment will not be applied to subjects over 18 years of age. No infants (< 12 months of age) will be enrolled, although subjects originally diagnosed with ALL during infancy are eligible if they are at least 12 months old at the time of study treatment initiation.

5.3 Adaptive Design

Not Applicable.

6. Covariates and Subgroups

6.1 Planned Covariates

Not Applicable.

6.2 Subgroups

Not Applicable

7. Definitions

Investigational Product (IP)

IP for this study refers to Carfilzomib.

Study Treatments

Study treatments include carfilzomib, chemotherapy backbone of R3 or VXLD, and IT chemotherapy drugs.

Study Day 1

The date of the first dose of carfilzomib or any chemotherapy components of R3 or VXLD.

Study Day

The number of days from the study day 1 to a date of interest, inclusive:

Study day = (date of interest – study day 1) + 1.

End of Study Date for a Subject

The end of study date is the date recorded on the End of Study page for an enrolled subject.

Duration of Treatment

Duration of treatment of a drug of interest is defined as the number of days = (date of last dose of the drug of interest – date of first dose of the drug of interest) +1. Duration of treatment with the chemotherapy combination is defined as the number of days = (date of last dose of any drug in the chemotherapy combination – date of first dose of any drug in the chemotherapy combination) +1.

Number of Actual Treatment Days

Number of actual treatment days is defined as the total number of days when a drug of interest was administered.

Total Dose

Total dose received of a drug of interest is defined as the cumulative total quantity that was actually administered for the specific drug of interest.

Average Dose

Average dose of a drug of interest is defined as the total dose received divided by the number of doses administered.

Actual Dose Intensity

Actual dose intensity of a drug of interest is defined as total dose of this drug divided by the product of baseline Body Surface Area (BSA) if applicable, and duration of treatment (days).

Relative Dose Intensity

Relative Dose Intensity (%) reflects whether the dose intensity of a drug of interest was implemented as planned. It is defined as 100 times the ratio of actual dose intensity to the intended dose intensity. Actual dose intensity is defined as above. Intended dose intensity is the planned cumulative dose divided by the protocol specified treatment duration in days. Specifically, planned cumulative dose is the sum of protocol specified

dose (details in Section 8.1.1, Section 8.1.2, Figure 2, Section 8.1.3, and Figure 3 in protocol).

Baseline Value

Baseline value is the closest of recorded measurements prior to or on the Study day 1. If there are multiple valid lab measurements for the same analyte collected on the same date and time, then the baseline will be decided conservatively. The change from baseline is defined as the value of interest minus baseline value.

Treatment emergent adverse events

Treatment emergent adverse events **(**TEAEs) are defined as adverse events (AEs) that start on or after the first administration of study drug and within the end of the study or 30 days of the last day of any study treatment**,** whichever occurs earlier. An AE that is present before the first administration of study treatment and subsequently worsens in severity during treatment is also considered to be treatment emergent.

DLT evaluable population

DLT evaluable population is defined as all subjects who received all planned dose of carfilzomib and the chemotherapy backbone per protocol during the Lead-in window and Induction Cycle, or received at least one dose of carfilzomib and the chemotherapy backbone per protocol and experienced a DLT prior to completion of the Lead-in Window or Induction Cycle, or as clinically indicated. The DLT evaluable status is collected separately on a log reviewed and approved by the study team.

Subject Incidence Rate

The subject incidence rate for a given event is defined as the number of subjects with one or more reported occurrence of the event divided by the number of subjects who had the opportunity to experience the event.

Primary refractory

Primary refractory is defined as failure to achieve a CR, CRp, or CRi within 42 days (inclusive) after the start of therapy following the initial diagnosis.

Refractory relapse

A subject with refractory relapse has received 1 or more cycles of anti-leukemia therapy (exclusive of Section 6.2 of the protocol) between the most recent relapse and the enrollment in this study, and failed to achieve a CR, CRp, or CRi. Number of refractory therapies is the number of therapies a subject has received and failed to achieve a CR,

CRp, or CRi, between the most recent relapse and the time they were enrolled in this study.

Duration of remission

Duration of remission is defined as the time from the date of a response to a prior therapy with a remission, i.e., a response at least CR, CRp, or CRi, to the date of relapse. (e.g., the time between the date of the first remission to the date of the first relapse is the duration of first remission, the time between the date of the second remission to the date of the second relapse is the duration of the second remission, etc).

Disease Response

Disease response will be assessed by investigator and includes the following categories: complete **remission** (CR), complete **remission** without platelet recovery (CRp), complete **remission** with incomplete hematologic recovery (CRi), partial response (PR), stable disease (SD), progressive disease (PD), induction death (ID) and not evaluable (NE). The criteria for determining treatment response are defined in Section 10.2 of the protocol.

Lead-in Window

The Lead-in Window consists of 1 week of carfilzomib administration as a single agent, during the DE 1 portion of the study. More details are presented in Section 8.1.1 of the protocol.

Induction Cycle

The Induction Cycle consists of a 4-week cycle of treatment that includes carfilzomib and a combination chemotherapy backbone. During the DE 1, the Induction Cycle begins on Day 8 of the carfilzomib single agent Lead-in-Window. During the DE 2, Induction is the first cycle of therapy that the subject will receive starting with Day 1. More details are presented in Section 8.1 of the protocol.

Consolidation Cycle

The Consolidation Cycle begins on day 36 of the Induction Cycle (7 days after the Day 29 response assessment of the Induction period), or when the absolute neutrophil count (ANC) is $>$ 750/mcl and platelet count $>$ 75000/mcl, whichever comes later. It consists of a single 4-week cycle of carfilzomib in combination with Children's Oncology Group (COG)- modified Berlin-Frankfurt-Münster (BFM) chemotherapy, and appropriate Central Nervous System (CNS) therapy. More details are presented in Section 8.1.3 of the protocol.

Subjects who reached SD, PR, CRp, CRi or CR at the end of their Induction Cycle may continue with this optional cycle.

8. Analysis Sets

8.1 Safety Analysis Set

The safety analysis set corresponds to the Safety Evaluable Population, defined as subjects receiving any amount of the study treatment regimen (chemotherapy backbone, carfilzomib, and any IT chemotherapy drugs).

Based on the treatment period that subjects have started during the trial, the safety analysis set will be further defined as:

- The Lead-in Window Safety Analysis set, which includes the safety evaluable subjects who started carfilzomib in the Lead-in Window
- The Induction Safety Analysis set, which includes the safety evaluable subjects who started the Induction cycle.
- The Consolidation Safety Analysis set, which includes the safety evaluable subjects who started the Consolidation Cycle.

8.2 Response-Evaluable Analysis Set

Response-evaluable analysis set corresponds to the Efficacy Evaluable Population, defined as the Safety Evaluable subjects who have a baseline disease assessment and at least one post-baseline disease assessment, or who dropped out due to AE prior to the first post-baseline disease assessment. Subsets of efficacy evaluable subjects will be further created based on the Safety Analyses subsets, as defined in Section **8**.1, as needed.

8.3 Interim Analyses Set(s)

Not Applicable.

9. Planned Analyses

9.1 Interim Analysis and Early Stopping Guidelines

No interim analysis is planned for this study.

9.2 Primary Analysis

No primary analysis is planned for this study.

9.3 Final Analysis

The trial will stop enrolling subjects when any of the stopping criteria specified in the protocol is satisfied. The Final analysis will be performed when the last subject enrolled had the opportunity to complete the study regimen and all required follow-up assessments.

10. Data Screening and Acceptance

10.1 General Principles

The objective of the data screening is to assess the quantity, quality, and statistical characteristics of the data relative to the requirements of the planned analyses.

10.2 Data Handling and Electronic Transfer of Data

The Amgen Global Study Operations-Data Management (GSO-DM) department will provide all data to be used in the planned analyses. This study will use the RAVE database. The database will be subject to edit checks outlined in the Clinical Data Management Plan (DMP).

10.3 Handling of Missing and Incomplete Data

Assumptions for missing or partially missing dates for important variables will be made to allow inclusion of appropriate data records in the analyses.

Imputed dates will not be presented in the listings. However, if the partial AE onset date information does not indicate whether the AE started prior to treatment or after the TEAE period, the AE will be classified as TEAEs.

The handling of incomplete and partial dates for adverse events, concomitant medications, and death are described in Appendix F1 and Appendix F3.

10.4 Detection of Bias

Not Applicable.

10.5 Outliers

Outlier data will not be excluded unless scientifically justified.

10.6 Distributional Characteristics

Not Applicable.

10.7 Validation of Statistical Analyses

Programs will be developed and maintained, while output will be verified in accordance with current risk-based quality control procedures.

Tables, figures, and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analyses consists of Amgen-supported versions of statistical analysis software. All statistical summaries and analyses will be performed in SAS System version 9.4 or later. The Bayesian algorithm is validated independently with EAST version 6 or higher.

11. Statistical Methods of Analysis

11.1 General Considerations

Summaries of all data will be presented by initial dose level cohorts for each dose escalation, separately. Unless otherwise specified, the data will be presented for Safety Analysis Set.

Summary statistics will be provided for selected endpoints. For continuous variables, the number of subjects with non-missing data (n), mean, standard deviation, median, minimum, and maximum will be presented. For discrete data, the frequency and percent distribution will be presented. Unless otherwise indicated, percentages will be calculated with respect to the number of subjects in the analysis population in each dose group.

Confidence intervals, when presented, will be constructed at the 2-sided 95% level. For binomial variables, exact distribution methods will be employed.

Individual subject listings will be produced for selected endpoints using data recorded on the electronic case report forms (eCRFs) or derived data.

11.2 Subject Accountability

The following subject disposition information will be summarized descriptively:

- **Number of enrolled subjects**
- Number of treated subjects (Safety population)
- Number (%) of subjects who discontinued investigational product and non-investigational products and the primary reasons for discontinuation. The percentage is calculated with respect to Safety population.
- Number (%) of subjects who discontinued study and the primary reason for discontinuation. The percentage is calculated with respect to Safety population.
- Number of subjects who have initiated Lead-in Window, Induction Cycle and Consolidation Cycle.

11.3 Important Protocol Deviations

Important Protocol Deviations (IPDs) categories are defined by the study team before the first subject's visit and updated during the IPD reviews throughout the study and prior to the database lock. These definitions of IPD categories, sub-category codes, and descriptions will be used during the study. The final IPD list provided by the study team will be used for the summary table and listing. Violations of the inclusion and exclusion criteria in the protocol will be included.

11.4 Demographic and Baseline Characteristics

11.4.1 Demographics

Demographic data includes age (continuous, categorical: < 2, 2-11, 12-17, > = 18, and EMA-specified age groups: New born 0-27 days, Infants and toddlers 28 days-23 months, Children 2-11 years, Adolescents 12-17 years, 18-64 years), sex, ethnicity, race, weight (kg), height (cm), BSA (m^2), and BMI (kg/m^2) at baseline.

11.4.2 Medical History

The number (%) of subjects who experienced a prior disease or disorder will be summarized by system organ class and preferred term.

11.4.3 Disease Characteristics

The following baseline disease characteristics will be summarized:

- Type of ALL
- Duration from initial diagnosis of ALL (in months)
- WBC result at diagnosis
- Primary refractory (a.k.a. Induction failure)
- Relapse
    - Refractory relapse
        - Number of refractory therapies after the qualifying relapse
- Number of relapses to any prior therapy
- Location(s) of the most recent relapse (bone marrow, CNS, testicular, and more as indicated by investigator)
- Number of prior therapies
- Duration of first remission (< 18 months, 18 to 36 months, and > 36 months)
- Duration of last remission prior to the qualifying relapse
- If received, days from prior allogeneic stem cell transplant to enrollment (days)
- Lansky Play-Performance Scale
- Peripheral Lymphoblasts Count
- Peripheral white blood cell count (WBC)
- Platelet counts
- Cytogenetics

The date of receiving prior stem cell transplant (SCT) is imputed as the stop date of the last drug in the corresponding regimen+2 days. The prior regimen corresponding to a SCT was determined as regimen containing total body radiation, busulfan, thio-tepa,

anti-thymocytes globulin, or specified as a stem cell transplant in ALL prior therapy history. The final list of subjects with prior SCT will be determined by clinical review.

11.5 Efficacy Analyses

The efficacy analyses will be based on subject data collected through study discontinuation or at the end of maximal treatment duration plus 30 days of safety follow-up, whichever occurs first. Although the primary analyses will use the Safety Analysis Set, additional efficacy analyses may be performed on the Response Evaluable Analysis Set.

Response assessment data including the response status and the MRD status ($< 10^{-3}$ and $< 10^{-4}$ lymphoblasts) will be listed.

Table 11.1. Efficacy Endpoint Summary Table

Endpoint	Primary Summary and Analysis Method (Specify Analysis Set if FAS is Not Used)	Sensitivity Analysis
Secondary Endpoints		
Combined proportion of subjects who achieve CR or CRp at the end of the Induction Cycle	• Summarize proportion of subjects who achieve CR or CRp at the end of the Induction Cycle. Generate 2-sided exact Clopper-Pearson 95% confidence interval. The analysis will be done in Induction Safety Analysis Set and Induction Response Evaluable Analysis Set.	NA
Proportion of subjects who achieve MRD status $< 10^{-3}$ and $< 10^{-4}$ lymphoblasts at the end of the Induction Cycle	• Summarize proportion of subjects who achieve MRD status $< 10^{-3}$ and $< 10^{-4}$ lymphoblasts at the end of the Induction Cycle for each assessment method. Generate 2-sided exact Clopper-Pearson 95% confidence interval. The analysis will be done in Induction Safety Analysis Set and Induction Response Evaluable Analysis Set. • Additionally, the proportion of subjects who do not have adequate data regarding the MRD status will be summarized, and the reasons for inadequate data will be tabulated.	NA

11.5.1 Analyses of Primary Efficacy Endpoint(s)

Not Applicable.

11.5.2 Analyses of Secondary Efficacy Endpoint(s)

The following analyses will be done on the induction safety analysis set and on the induction response evaluable set.

Summarize proportion of subjects who achieve CR or CRp at the Induction Cycle Day 29. The 2-sided exact Clopper-Pearson 95% confidence interval will be generated. The analysis will be done in Induction Safety Analysis Set and Induction Response Evaluable Analysis Set.

Summarize proportion of subjects who achieve MRD status $< 10^{-3}$ and $< 10^{-4}$ lymphoblasts at the Induction Cycle Day 29. Generate 2-sided exact Clopper-Pearson 95% confidence interval. The analysis will be done in Induction Safety Analysis Set and Induction Response Evaluable Analysis Set. Additionally, the

proportion of subjects who do not have adequate data regarding the MRD status will be summarized, and the reasons for inadequate data will be tabulated.

11.5.3 Analyses of Exploratory Efficacy Endpoint(s)

[redacted]

11.6 Safety Analyses

The safety analyses will be based on the safety evaluable population and will be presented separately for each dose escalation and each dose level. All summaries will be presented by the assigned dose level for each safety analysis subset as defined in Section **8**.1.

11.6.1 Analyses of Primary Safety Endpoint(s)

The safety and tolerability will be assessed through summaries of study drug administration, review of the incidence of TEAEs, SAEs, selected search strategies for specified medical concepts, changes from baseline in selected laboratory analytes, vital signs, and physical findings. These primary safety endpoints will be summarized separately for the carfilzomib single-agent Lead-in Window and the Induction Cycle treatment periods.

For Lead-in window safety analysis set, TEAEs that started within the first dose of Lead-in window treatment (inclusive) to the earlier of the end of safety follow-up (inclusive), or the first dose of induction cycle treatment (not inclusive), will be included in this analysis.

For the Induction safety analysis set, TEAEs that started within the first dose of induction cycle treatment (inclusive) to the earlier of the end of safety follow-up (inclusive), or the first dose of consolidation cycle treatment (not inclusive) will be included in this analysis.

11.6.2 Analyses of Exploratory Safety Endpoints

[redacted]

11.6.3 Adverse Events

Each reported AE term will be mapped to a preferred term (PT) and a system organ class (SOC) using the Medical Dictionary for Regulatory Activities (MedDRA, version 23.1).

The subject incidence of TEAEs will be summarized for all TEAES, TEAEs of grade 3 or higher, treatment-related AEs, serious AEs, AEs leading to discontinuation of investigational product, fatal AEs, tabulated by SOC and PT in descending order of frequency and by PT in descending order of frequency without regard to SOC. The maximum NCI - CTCAE 4.03 toxicity grade and strongest causal relationship to study treatment will **be** evaluated, and the TEAEs will be also summarized by severity and relationship to the study drugs.

Subject incidence of events of interest (EOI), grade 3 or higher EOI, and serious EOI as standardized MedDRA queries and/or Amgen customized queries will be summarized by PT in descending order of frequency.

An overall summary of TEAEs will present the number (%) of subjects with:

- with at least one TEAE
- with at least one treatment-related TEAE
- with at least one carfilzomib-related TEAE
- with at least one grade > = 3, = 5 TEAE
- with at least one treatment-related grade > = 3, = 5 TEAE
- with at least one SAE
- with at least one grade 3 or higher hematological AEs
- with TEAE leading to discontinuation of all study drugs
- death within 30 days of last dose of study drug

Hematological AEs includes the Blood and Lymphatic system disorders SOC and the Haematology investigations (including blood groups) HLGT.

The DLTs are assessed by investigators according to the definition described in Section 8.3 of the protocol.

The subjects who experienced DLT, the detail of subject's DLT(s), and subjects' DLT evaluable status will be summarized in subject level listings.

All TEAE data will be listed by study site, dose level, subject identification number, and study day. Similar listings will be produced separately for all SAEs.

11.6.4 Laboratory Test Results

Actual values and change from baseline for laboratory measurements will be summarized.

The hematology laboratory analytes of interest are:

- CBC with differential (i.e., Total Neutrophils, Eosinophil, Basophils, Lymphocytes, Monocytes, Lymphoblasts) and platelet count

The chemistry laboratory analytes of interest are:

- Electrolytes: Sodium, Potassium, Chloride, Bicarbonate, Corrected Calcium, Magnesium, Phosphorus, Phosphate
- Blood glucose
- Renal function: blood urea nitrogen (BUN), Creatinine
- Hepatic function: Aspartate Aminotransferase (AST), Alanine Aminotransferase (ALT), total and direct bilirubin, Alkaline Phosphatase (ALP), albumin
- Pancreatic function: amylase, lipase
- Uric acid
- Bone-specific ALP

All laboratory data will be summarized by NCI-CTCAE grade. Shift of grade in laboratory test results will be summarized.

11.6.5 Vital Signs

Actual value and change from baseline for vital sign results including blood pressure, pulse, respiratory rate, and temperature will be summarized.

11.6.6 Physical Measurements

Actual value and change from baseline for the physical measurements, including height, weight, and BSA, will be summarized.

11.6.7 Electrocardiogram

The electrocardiogram (ECG) measurements from this clinical study were performed as per standard of care for routine safety monitoring and will be presented in subject listings.

11.6.8 Antibody Formation

Not Applicable.

11.6.9 Exposure to Investigational Product

Descriptive statistics in terms of treatment duration (days), number of actual treatment days, total dose received, number of dose administered, average dose per administration, actual dose intensity, and relative dose intensity will be produced to describe the exposure to carfilzomib.

11.6.10 Exposure to Other Protocol-specified Treatment

All the analyses planned for IP will be carried over for each drug in any chemotherapy combination and CNS treatment combination, if applicable. For CNS combination drugs, exposure will be summarized for CNS positive and negative separately. Number of oral administration and total dose received will be summarized for trimethoprim-sulfamethoxazole (TMP-SMX).

11.6.11 Exposure to Concomitant Medication

The number (%) of subjects receiving concomitant medication will be summarized by medication PT as coded by the World Health Organization Drug (WHO DRUG) dictionary version September 2018 or later in descending order of frequency.

11.7 Other Analyses

[redacted]

11.7.1 Analyses of Pharmacokinetic or Pharmacokinetic/ [redacted] Endpoints.

The PK parameter estimates for carfilzomib will be estimated from plasma concentration-time profiles using standard noncompartmental approaches over the complete sampling interval. Carfilzomib PK parameters may include area under the concentration-time curve (AUC), maximum observed plasma concentration (C_{max}), and if feasible, plasma terminal half-life. Parameter estimates will be tabulated and summarized (i.e., mean, standard deviation). These data will also be used in a population PK analysis to evaluate the effects of subject demographic characteristics and covariates on carfilzomib PK parameters. [redacted]

[redacted]

PK [redacted] analyses will be provided by Clinical Pharmacology, Modeling & Simulation (CPMS). Details of this analysis will be reported separately.

[redacted] Details of this analysis will be reported separately.

11.7.2 Analyses of Clinical Outcome Assessments

Not Applicable.

11.7.3 Analyses of Health Economic Endpoints

Not Applicable.

[redacted]

12. Changes from Protocol-specified Analyses

In protocol Section 13.6.4, it is specified that "For the proportion of subjects with a decreased percentage of lymphoblasts in the bone marrow at the end of a 1-week carfilzomib single-agent Lead-in Window, the 1-sided exact Clopper-Pearson 95% confidence interval will be generated". In this SAP, the confidence interval was changed to 2-sided.

In protocol Section 13.6.2, it is specified that “Relevant laboratory and vital sign (temperature, heart rate, respiratory rate, and blood pressure) data will be displayed by visit and time (when available), with NCI-CTCAE (Version 4.03) Grade 3 and 4 values identified where appropriate.” In this SAP, vital signs will not be graded by NCI-CTCAE (version 4.03) and not displayed with NCI-CTCAE (version 4.03) grade 3 and 4 values.

In protocol Section 13.1.1, the primary endpoint is specified that ‘Determination of the MTD as the dose that has the highest posterior probability of having a DLT rate within the target toxicity interval (20%–33%), while the posterior probability of excessive/unacceptable toxicity (> 33%–100%) is less than 40% (per BLRM algorithm), or the stopping rules as specified in Section 5.2, and per CSRC recommendation.’ In this SAP, the endpoint for the primary objective of determination of MTD is incidence of DLT.

13. Literature Citations / References

Neuenschwander B, Branson M, Gsponer T. Critical aspects of the Bayesian approach to phase I cancer trials. Statist Med. 2008; 27:2420–39.

14. Appendices

Appendix A. Reference Values/Toxicity Grades

Laboratory Values

Safety laboratory values below a distinct limit (e.g., detection limit documented as "< [limit]") will be substituted by half of the limit and values above a distinct limit (documented as "> [limit]") will be substituted by the limit itself for all analyses. A Grade will be assigned to each laboratory based on CTCAE version 4.03. Values not meeting any of the criteria will be assigned a grade 0.

Appendix B. Phase 2 Statistical Analysis Plan

15. Introduction

This part of the document includes the details of the Statistical Analysis Plan (SAP) that have been outlined within the phase 2 protocol **amendment 12** for study 20140106, dated **28 August 2023**. **The scope of this plan includes the interim futility analysis, the primary analysis and the final analysis that are planned and will be executed by the Amgen Global Biostatistical Science department unless otherwise specified.**

16. Phase 2 Objectives, Endpoints and Hypotheses

16.1 Objectives and Endpoints

All primary and secondary objectives will be analyzed for T-cell and B-cell **acute lymphoblastic leukemia (**ALL) populations separately, unless otherwise specified. **For the primary and secondary endpoint evaluation, treatment response will be derived by the sponsor based upon local evaluation of bone marrow, peripheral blood, and extramedullary disease status. The definitions of complete remission (CR), complete remission without platelet recovery (CRp), complete remission with partial hematological recovery (CRh), complete remission with incomplete hematologic recovery (CRi), stable disease (SD) and progressive disease (PD) are described in Appendix G according to the criteria in Table 26 of protocol. Minimal residual disease (MRD) status will be determined per central lab review of bone marrow MRD using NGS.**

Objectives	Endpoints
Primary	
• Compare the rate of CR of CFZ-VXLD at the end of induction therapy to an appropriate external control.	• CR after induction therapy
Secondary	
• Evaluate the safety and tolerability of CFZ-VXLD	• Incidence of treatment-emergent and treatment-related adverse events and severe adverse events and laboratory abnormalities during the induction therapy and consolidation therapy
• Compare the rate of **CR,** CRp, CRh, and CRi of CFZ-VXLD at the end of induction therapy relative to an appropriate external control	• **CR,** CRp, CRh and CRi at the end of induction therapy
• Compare EFS for CFZ-VXLD to an appropriate external control	• EFS, defined as time from initiation of therapy until treatment failure (defined as failure to reach at least a CRi after consolidation or after induction in subjects

	that do not receive consolidation), relapse, or death from any cause
• Compare OS for CFZ-VXLD relative to an appropriate external control	• OS defined as time from initiation of therapy until death from any cause
• Estimate the DOR for CFZ-VXLD and an appropriate external control	• DOR, defined a**s** time from **earliest of** CR, CRp, CRh, or CRi to relapse or death from any cause
• Estimate the rate of MRD[-] at the end of induction in subjects receiving CFZ-VXLD	• MRD **status** using NGS **less than** 10^{-4} after induction therapy in subjects achieving CR
• Estimate the rate of (MRD[-]) bone marrow after induction and consolidation therapy in subjects with B-cell ALL **or T-cell ALL** receiving CFZ-VXLD	• MRD **status** using NGS **less than** 10^{-3} and **less than** 10^{-4} in subjects achieving CR, CRp, CRh, or CRi after induction and consolidation therapy, separately
• Estimate the proportion of subjects that bridge to stem cell transplant or CAR-T cell therapy in subjects receiving CFZ-VXLD	• Occurrence of a stem cell transplant or CAR-T, without an intervening relapse after protocol specified therapy
• Estimate the rate of CR, CRp, CRh and CRi of CFZ-VXLD at the end of consolidation therapy in subjects receiving CFZ-VXLD	• CR, CRp, CRh, and CRi after consolidation therapy
• Estimate the pharmacokinetics of carfilzomib when administered as part of VXLD regimen	• Carfilzomib pharmacokinetics parameters, including AUC, C_{max}, and if feasible $t_{1/2}$
Exploratory	
[redacted]	

16.2 Hypotheses and/or Estimations

The primary hypothesis is that CFZ-VXLD is superior to standard-of-care with regards to CR rate. This corresponds to evaluating whether the 95% CI for the **adjusted** odds ratio

(OR) between CR rate of CFZ-VXLD subjects versus external control subjects excludes 1 for each phenotype (B-cell/T-cell) independently **(H_0: lower bound of 95% CI for OR is less than or equal to 1; H_1: lower bound of 95% CI for OR is higher than 1)**.

17. Study Overview

17.1 Study Design

The phase 2 portion of the study is a multicenter, single-group, externally-controlled study of carfilzomib in combination with VXLD in a minimum of 100 pediatric subjects, unless futility is met. Eligible subjects must be **greater than or equal to** 1 month to **less than** 21 years old, with their original diagnosis at **less than** 18 years of age, and must have ALL with bone marrow relapse (**greater than or equal to** 5% leukemia blasts in bone marrow) or refractory relapse with or without extramedullary disease of the T-cell phenotype or of the B-cell phenotype after having received a targeted B-cell immune therapy (e.g., blinatumomab, inotuzumab, **or** CAR-T therapy). Infants enrolling on this study will receive a modified VXLD and optional consolidation based on the dose and schedule used in COG AALL15P1 study.

The carfilzomib dose for phase 2 will be 20 mg/m^2 on day 1 of induction, 56 mg/m^2 on days 2, 8, 9, 15, and 16 of induction and days 1, 2, 8, 9, 15, and 16 of consolidation.

Eligible subjects will be treated with VXLD plus carfilzomib for 1 cycle (induction), followed by assessment of treatment response. All subjects who do not show progression during induction will undergo a bone marrow and extramedullary disease evaluation after completion of induction therapy between day 29 to 45 of induction (between day 35 to 50 for infants), based on blood count recovery but before the start of **post-induction** therapy, whichever comes first. Response will be assessed per **local and** central laboratory review of bone marrow, peripheral blood **and differential**, and local assessment for sites of extramedullary disease. MRD will be assessed by NGS central laboratory review and local evaluation by flow cytometry and/or PCR or NGS when available.

Subjects without disease progression **(as defined in protocol Table 26)** after induction may, at the investigator's discretion, be treated with 1 cycle of consolidation chemotherapy plus carfilzomib. Subjects who do not show disease progression during consolidation will be assessed for treatment response between day 29 to 45 of consolidation (between day 36 to 50 for infants), based on blood count recovery prior to the start of alternative therapy.

Treatment response after consolidation therapy will also be assessed per **local and** central laboratory review of bone marrow, peripheral blood counts and differential, and local assessment for sites of extramedullary disease. **MRD will be assessed by NGS central laboratory review and local evaluation by flow cytometry and/or PCR or NGS when available**. The consolidation treatment will consist of one 28-day cycle of Children's Oncology group-modified Berlin-Frankfurt-Munster (BFM) therapy **as detailed in the protocol**. **After completion of study therapy, subjects will be followed for subsequent treatment(s), event-free survival, overall survival, duration of remission and ability to bridge to transplant.**

For selected objectives, the treatment response to induction of subjects receiving CFZ-VXLD will be compared to an external control arm of subjects from an observational study of relapse**d** or refractory pediatric ALL (study 20180065) that received standard-of-care (SoC) therapy, after appropriate adjustment using logistic regression model with inverse probability treatment weights (IPTW) for the average treatment effect of the treated (ATT). These weights are derived for each CFZ-VXLD and control subject based on their propensity score (PS) for receiving CFZ-VXLD. More details provided in Section 2**3**.**6**.1 [redacted]

17.2 Sample Size

The sample size was determined based on practical considerations and limited **phase 1** data available as per Section **18**.2.4 of the protocol. Due to the rarity of the subject population, **at least 100 subjects are expected to receive at least one dose of carfilzomib in phase 2 part of the study. A minimum of 50** B-cell **ALL subjects** and **a minimum of 30** T-cell ALL subjects are planned to be enrolled, in order to reach more than 70% power for testing the primary hypothesis for at least one phenotype (B-cell or/ and T-cell) at 2**-**sided alpha of 0.05 without multiplicity adjustment, with a **non-binding interim** futility analysis.

The external control arm **will be selected from the observational study 20180065 which retrospectively collects data on pediatric patients with relapsed and refractory T-cell and B-cell ALL treated with contemporary salvage therapy. The external control arm will be selected to be as similar as possible to subjects enrolled in Study 20140106, as described in study 20180065 protocol/statistical analysis plan (SAP).**

The external control arm is expected to include **approximately 74** B-cell ALL subjects and approximately **60** T-cell ALL subjects **as per the Primary Analysis Sets (T-PAS/B-PAS) included in Study 20180065 protocol/SAP**.

For selected objectives, the response and disease status for each subject will be derived by the sponsor based upon the local assessments and following the definitions included in protocol Table 26. The details of the derivation algorithm are specified in the Appendix G.

Additionally, based on the characteristics of subjects enrolled into the phase 1b part of the study, and after appropriate propensity score adjustment, the CR rates in the external control arms are expected to be 25% for B-cell and 15% for T-cell subjects **respectively**, and approximately 60% for B-cell and 30% for T-cell subjects from CFZ-VXLD arm. As these estimates are based on data available so far, **as detailed** in Section **18**.2.4 of the study 20140106 protocol, additional scenarios proposed for experimental arm **CR rates, sample sizes** and the power**/external control sample sizes** expected in these cases are included in Table **17.1**.

Table 17.1 Sample Size and Power Estimates for Hypothesized Treatment Effect

Phenotype	Control Arm CRR	CFZ-VXLD Arm CRR	Control Arm Sample Size	CFZ-VXLD Arm Sample Size	Power*
B-cell	0.25	0.5	74	50	0.70
B-cell	0.25	0.55	74	50	0.85
B-cell	0.25	0.6	74	50	0.94
B-cell	0.25	0.5	66	56	0.72
B-cell	0.25	0.55	66	56	0.86
B-cell	0.25	0.6	66	56	0.94
B-cell	0.25	0.5	62	60	0.71
B-cell	0.25	0.55	62	60	0.87
B-cell	0.25	0.6	62	60	0.95
T-cell	0.15	0.3	60	30	0.24
T-cell	0.15	0.3	60	40	0.27

[a] The power was adjusted by a factor of 0.9**9** due to the **interim analysis for** futility.
CRR= complete remission rate

The hypothesized treatment effect size was determined based on the covariates predictive of CR observed in the phase 1 portion of the trial and the CR rates reported in

the literature for those characteristics (more details **are** provided in protocol Section **18**.2.4).

For each phenotype separately (B-cell/T-cell), H_0 will be rejected at the primary analysis if the 95% confidence interval (CI) **of the adjusted OR** from the propensity adjusted model is above 1. The power was estimated assuming that both treatment arms will be balanced with respect to baseline prognostic factors after propensity score adjustment, the control arm **probability of CR (**PoCR**)** will be consistent at the time of interim futility analysis and primary analysis, the odds ratio of PoCR is 1 under H_0, and the **power loss is at most** 0.0**1** under H_1. The power was adjusted by a factor of 0.9**9** due to futility **check**.

Since the variance of the log odds ratio will be inflated with the IPTW-ATT method compared to a non-weighted model, the estimates were adjusted by a variance inflation factor of 1.4 obtained via simulation using data from the blinatumomab B-cell ALL pediatric US study 20140228 and EU study 20120299. The inflation factor was selected as the empirical 95% confidence interval upper limit using the sandwich estimator of the log odds ratio standard error. A grid search approach was used to estimate the sample size needed for the test. For each candidate number of subjects for the test arm, the lower bound of the 95% CI of the odds ratio (incorporating variance inflation factor) for each possible pair of outcomes was calculated and compared with 1. Power was derived as the summation of the joint probabilities of those pairs of outcomes with a lower bound greater than 1. T[redacted]

The final selection of external control arm subjects **included in the Primary Analysis Set (PAS)** is anticipated to occur prior to the **primary** analysis, but not prior to initiation of study enrollment. The interim futility analysis **set** will include **approximately 6**0 subjects **(60% of the planned number of subjects) from the experimental CFZ-VXLD arm who have received at least one dose of carfilzomib** and **have had the opportunity to complete a post induction response evaluation, and all available subjects from the external control arm by the time of interim analysis (as defined in Interim Analysis Set [T-IAS/B-IAS] of Study 20180065 protocol/SAP)**.

18. Covariates and Subgroups

18.1 Planned Covariates

[redacted]

18.2 Subgroups

Primary **endpoint** (CR) and selected secondary endpoints (CRi or better, EFS, **and** OS) will be explored in the following subgroups of interest using the Primary Analysis **S**et **(as defined in Section 20.1.1)**. Analyses will be performed separately for B-cell and T-cell ALL populations, unless otherwise specified. When the number of subjects in the subgroup is not sufficient (i.e., less than 20% of subjects in a treatment arm), relevant subgroups may be combined. In these subgroup analyses, all treatment arm subjects will receive a weight of 1 while control arm subjects will be IPTW-ATT weighted, with the same weights used for the primary analysis of the primary endpoint. Odds Ratio (OR) and 95% CI will be produced for CR, and CRi or better. Hazard Ratio (HR) and 95% CI will be produced for EFS and OS.

- History of prior **allogeneic** HSCT (yes, no)
- **History of primary induction failure (yes vs no)**
- Relapse time from **prior allogeneic** HSCT **(months)** [**<12, ≥12, unknown (including patients without HSCT)]**
- Early T cell precursor phenotype at time of initial diagnosis (yes, no, unknown) (for T cell phenotype only)
- Number of prior relapses (1, 2, > 2)
- Refractory **qualifying** relapse (yes, no)
- Cytogenetic **risk at time of initial diagnosis (high risk, non-high risk, unknown**. High risk = t [9:22], MLL translocation, hypodiploidy [< 44 chromosomes]) (only for B-cell phenotype)
- Duration of first remission **(months)** (< 18, 18 to 36, > 36, **missing**)
- Age at **diagnosis (years)** (**<1**, **1** to < 2, 2 to < 10, ≥ 10)
- **Sex (male, female, unknown)**
- **Isolated bone marrow at relapse (yes, no)**
- **Blast percent in bone marrow prior to the qualifying treatment (<50%, ≥50%)**
- Geographical region (**Region 1, Region 2**)
  - **Region 1 includes countries that participated in 20140106 phase 1 (US, Denmark, France, Italy, Spain, Austria, Australia)**

- **Region 2 includes all other countries added to 20140106 phase 2 that did not participate in phase 1**

The additional secondary endpoints: occurrence of a SCT or CAR-T, and MRD [-] status will be estimated in the same subgroups listed above, using the Safety Analysis set with no adjustment.

19. Definitions

Average Dose

Average dose of a drug of interest is defined as the total dose received divided by the number of doses administered.

Actual Dose Intensity

Actual dose intensity of a drug of interest is defined as total dose of this drug divided by the product of baseline Body Surface Area (BSA) if applicable, and **the number of actual treatment days**. **The days without dosing during the treatment period will not be included.** For BSA calculation rule, please refer to protocol.

Relative Dose Intensity

Relative Dose Intensity (%) **of a drug of interest** reflects whether the dose intensity of **the specified** drug was implemented as planned. It is defined as 100 times the ratio of actual dose intensity to the intended dose intensity. Intended dose intensity is defined as the planned cumulative of study treatment **per the** schedule listed in protocol divided by the **number of the** protocol specified treatment days.

Number of Actual Treatment Days

Number of actual treatment days **of a drug of interest** is defined as the total number of days when a drug of interest was administered.

Baseline Value

Baseline value is the closest of recorded measurements prior to or on the Study day 1. The change from baseline is defined as the value of interest **minus** baseline value.

Cytogenetics Risk

The high risk cytogenetics for subjects in the B cell ALL group is defined by the presence of t (9:22), MLL translocation, or hypodiploidy (<44 chromosomes). **The non-high risk group consists of subjects with known cytogenetics data that is none of the above three cytogenetic findings.**

Complete remission (CR) rate

The CR rate is the proportion of subjects whose best response is CR at the end of induction therapy.

Rate of CRi or better

The rate of CRi or better will be used to summarize the secondary endpoints of CR, CRp, CRh, and CRi (i.e., CRi or better) after induction therapy and consolidation therapy, separately.

- **The rate of CRi or better after induction is defined as the proportion of subjects whose best response after induction therapy and prior to consolidation is CRi, CRh, CRp or CR in Induction Safety Analysis Set.**
- **The rate of CRi or better after consolidation is defined as the proportion of subjects in Consolidation Safety Analysis Set whose best response after consolidation therapy is CRi, CRh, CRp or CR.**

MRD status using NGS less than 10^{-4} after induction therapy in subjects achieving CR

It is defined as the proportion of subjects in Induction Safety Analysis Set with achievement of CR and MRD negativity as assessed by NGS method at a 10^{-4} threshold after induction therapy.

MRD status using NGS less than 10^{-3} and less than 10^{-4} in subjects achieving CRi CRi, CRh, CRp, or CR after induction and consolidation therapy

It is defined as the proportion of subjects in CFZ-VXLD arm with achievement of CRi or better and MRD negativity as assessed by NGS method at 10^{-3} and 10^{-4} threshold after induction therapy (in Induction Safety Analysis Set) and consolidation therapy (in Consolidation Safety Analysis Set) separately.

MRD status less than 10^{-3} and less than 10^{-4} by flow cytometry and/or PCR or NGS in subjects achieving CRi, CRh, CRp, or CR after induction and consolidation therapy

It is defined as the proportion of subjects in CFZ-VXLD arm with achievement of CRi or better and MRD negativity as assessed by flow cytometry and/or PCR or NGS method (if one assessment method used, or, if MRD negative in all assessments; if MRD detectable in any method used subject not considered MRD negative) at 10^{-3} and 10^{-4} threshold performed by local site after induction therapy

(in Induction Safety Analysis Set) and after consolidation therapy (in Consolidation Safety Analysis Set) separately.

Duration of Treatment

Duration of treatment of a drug of interest (days) = (date of last dose of the drug of interest – date of first dose of the drug of interest) +1.

Duration of treatment with the chemotherapy combination (days) = (date of last dose of any drug in the chemotherapy combination – date of first dose of any drug in the chemotherapy combination) +1.

Duration of **Remission** (DOR)

Duration of **remission** is defined as the number of days between the date of the first tumor assessment indicating **CRi or better (CRi, CRh, CRp or CR)** through to the subsequent date of **relapse/**progression or death due to any cause, or where applicable date of censoring. Subjects who reached **CRi or better** and have not relapsed nor died by the analysis date will be censored at last **known alive** date. Subjects who do not achieve **CRi or better** will be excluded from the analysis of duration of **remission**.

DOR (days) = relapse/death date or censoring date – the first date of CRi or better + 1

Event-free survival (EFS)

EFS will be measured from the initiation of therapy until treatment-failure (defined as failure to reach CRi or better after consolidation **therapy** or after induction **therapy** in subjects that do not receive consolidation), relapse or death to any cause, **whichever occurred first. The treatment failure date is the date of PD or SD or non-evaluable (NE) per sponsor's response assessment. The relapse date is captured on the eCRF form Relapse Status & Subsequent Anti-Cancer.** Subjects **who are** alive and event free **are censored** at last **known alive date through the DCO date**.

EFS (months) = (treatment failure/relapse/death date or censoring date – first dose date of IP + 1) / 30.4

End of Study Date

The end of study date is defined as the date when the last surviving subject completes the safety follow-up, discontinues the study due to death, is lost to follow-up, withdraws full consent from the study, or the sponsor in consultation with regulatory authorities decides to close the study, whichever occurs first.

History of primary induction failure

Failure to achieve a response (i.e., CR/CRp/CRh/CRi) within 16 days after completion of the initial induction treatment following the subject's diagnosis of ALL.

Investigational Product (IP)

IP for this study refers to carfilzomib.

Last Known Alive Date

Last known alive date is the latest date before the death date during the study, according to the dates recorded on eCRFs and the dates in the data collected by the vendors.

Overall Survival (OS)

Overall survival is defined as the time in months from initiation of therapy until death due to any cause through the analysis data cutoff (DCO) date. OS is censored at the last **known alive** date through the DCO date. Data available after the DCO date will not be used to derive overall survival. Subjects with partial dates will be imputed according to **Section 22.3 and Appendix F**.

OS (months) = (death date or censoring date – first dose date of IP + 1) / 30.4

Study Day 1

Study day 1 corresponds to the date of **first dose of** carfilzomib.

Study Day

Study day = (date of interest – study day 1) + 1.

Study Treatments

Study treatments **(including protocol approved alternative agents)** include carfilzomib, chemotherapy backbone of VXLD**, modified BFM therapy** and IT chemotherapy drugs.

Study Treatment	**Induction Therapy**	**Consolidation Therapy**
Carfilzomib	✓	✓
Vincristine (VXLD, Modified BFM)	✓	✓
Dexamethasone (VXLD) (alternatives: Prednisone, Prednisolone, Methylprednisolone)	✓	
PEG-asparaginase (VXLD, Modified BFM) (alternatives: L-asparaginase, Erwinia, Rylaze)	✓	✓

Daunorubicin (VXLD) **(alternatives: Doxorubicin, Epirubicin, Idarubicin)**	✓	
Cytarabine (Modified BFM)		✓
6-Mercaptopurine (Modified BFM)		✓
Cyclophosphamide (Modified BFM)		✓
Methotrexate (IT chemotherapy)	✓	✓
Cytarabine (IT chemotherapy)	✓	✓
Hydrocortisone (IT chemotherapy) **(alternatives: other steroids)**	✓	✓

Subject Incidence

The subject incidence for a given event is defined as the number of subjects with one or more reported occurrence of the event divided by the number of subjects who had the opportunity to experience the event.

Total Dose

Total dose received of a drug of interest is defined as the cumulative total quantity that was actually administered for the specific drug of interest.

Treatment Emergent Adverse Event

Treatment-emergent adverse events (TEAEs) are defined as Adverse Events (AEs) starting on or after the first administration of study drug and within the end of the study or 30 days of the last day of any study treatment, whichever is earlier.

Treatment Response

For the main analysis of primary and selected secondary endpoints, treatment response will be **derived by sponsor based upon local evaluation of bone marrow, peripheral blood, and extramedullary disease status. The definitions of the following response categories are described in Appendix G according to the criteria included in Table 26 of protocol amendment 12:** complete **remission** (CR), complete **remission** without platelet recovery (CRp), complete **remission** with partial hematological recovery (CRh), complete **remission** with incomplete hematologic recovery (CRi), stable disease (SD), progressive disease (PD), and non-evaluable (NE). The criteria for determining treatment response for study 20140106 are consistent with the criteria used in study 20180065 protocol as detailed in Section 9.3.2.

20. Analysis Sets

Unless otherwise specified, all analysis sets below will be analyzed for B-cell **and** T-cell phenotype**s** separately.

20.1 Primary Analysis Set (PAS)

The primary analysis set will include subjects enrolled in the experimental CFZ-VXLD arm who received at least 1 dose of carfilzomib and all **external control** subjects **included** in the **T-PAS/B-PAS** as defined **in** Study 20180065 **protocol/SAP**. If **external control** subjects have multiple qualifying therapies, the **last qualifying** therapy will be chosen.

20.2 Safety Analysis Set

This analysis set will include subjects enrolled to the experimental CFZ-VXLD arm who received at least 1 dose of carfilzomib.

The following two analysis sets will be further defined based on the treatment period that subjects have started during the trial.

- **Induction Safety Analysis Set includes subjects who started the Induction cycle and received at least 1 dose of carfilzomib during the induction period.**
- **Consolidation Safety Analysis Set includes subjects who started the consolidation cycle and received at least 1 dose of carfilzomib during the consolidation period.**

20.3 Per Protocol Set(s)

This analysis set will include subjects enrolled to the experimental CFZ-VXLD arm who receive at least one dose of carfilzomib and **do not have any major** protocol deviations that might impact the efficacy endpoints. **Subjects with the following important protocol deviation (IPD) will be excluded from the per protocol set**. **The IPD numbers and descriptions are documented in the study level IPD list document.**

- **Major inclusion criteria related IPD: 111, 112, 113, 114, 118, 119**
- **Major exclusion criteria related IPD: 211, 214, 215, 216, 219, 220, 222, 228**
- **Major treatment non-compliance related IPD: 401, 403, 908**

20.4 Interim Futility Analyses Set(s)

The interim futility analysis set will include **approximately 60 subjects** (**6**0% of the planned number of subjects) from the experimental CFZ-VXLD arm who have received

at least one dose of CFZ and have had the opportunity to complete a post induction response evaluation **and all available subjects from the external control arm by the time of Interim Analysis (as defined in Interim Analysis Set [T-IAS/ B-IAS] of Study 20180065 protocol/SAP).**

20.5 VXLD/ VPLD Analysis Set (VAS)

This analysis set will include all subjects in **experimental** CFZ-VXLD arm who have received at least 1 dose of CFZ, while the external control arm includes only subjects who have been treated with VXLD/VPLD **as defined in Study 20180065 protocol/SAP**.

20.6 Chemotherapy Analysis Set (CAS)

This analysis set will include all subjects in **experimental** CFZ-VXLD arm **who received at least 1 dose of carfilzomib**, while the external control arm includes only subjects who have been treated with any ALL chemotherapy regimen, including proteasome inhibitors other than carfilzomib **as defined in Study 20180065 protocol/SAP.**

21. Planned Analyses

For each of the following planned analysis, data will be subject to ongoing checks for integrity, completeness, and accuracy in accordance with the Data Management Plan with the expectation that all outstanding data issues are resolved ahead of the snapshot. The data will be locked to prevent further changes, and a snapshot of the locked database will be used in the analysis.

21.1 Interim Analysis and Early Stopping Guidelines

An interim analysis for non-binding futility assessment in terms of the primary endpoint (CR after induction therapy) will be conducted independently for each phenotype (T-cell and B-cell) on the Interim Futility Analysis Set. The interim futility analysis will occur after at least 60 subjects in the experimental CFZ-VXLD arm have received at least one dose of carfilzomib and had the opportunity to complete a post induction response evaluation. The futility criterion will be derived such that the probability of futility is at least 70% under H_0 and the power loss is at most 0.01 under H_1, using a Bayesian predictive probability approach. Additional details for the interim futility assessment are described [redacted]

An Independent Biostatistics Group (IBG) will perform the interim futility analyses **according to the steps described** [redacted] and provide the **results as part of a closed session** report to an independent Data Monitoring Committee (DMC). To minimize the potential introduction of bias to the conduct of the study, members of the

DMC **and** IBG will not have any direct contact with study site personnel or subjects. The DMC will communicate major safety concerns and recommendations regarding study modification or termination based on **a risk/ benefit assessment using** the safety and efficacy **summaries produced by IBG** in accordance with the DMC charter.

Over the course of the study, DMC will review all available safety **and efficacy** data every 6 months **using as-is cumulative snapshots. The IBG will generate and provide routine open-session and closed-session reports to the DMC**.

Records of all meetings will be maintained by the IBG and DMC for the duration of the study. These records will be transferred and stored in the eTMF (in accordance with SOP-427356) after the termination of DMC meetings. Further details are provided in the DMC charter.

21.2 Primary Analysis

The primary analysis for each phenotype will occur when the external control subjects have been selected **from T-PAS/B-PAS** and **at least 100 enrolled** subjects in the experimental arm **have received at least one dose of carfilzomib and** had the opportunity to **complete** a post induction response evaluation. All the available data on primary, secondary and exploratory endpoints will be summarized at this time.

21.3 Final Analysis

The final analysis will occur **after** the end of the study, and it will be performed on all subjects included in the **Primary** Analysis Set.

22. Data Screening and Acceptance

22.1 General Principles

The objective of the data screening is to assess the quantity, quality, and statistical characteristics of the data relative to the requirements of the planned analyses.

22.2 Data Handling and Electronic Transfer of Data

The Amgen Global Study Operations-Data Management (GSO-DM) department will provide all data of the experimental CFZ-VXLD arm to be used in the planned analyses. This study will use the RAVE database. Data from the observational study will be collected and provided to Amgen by TACL as specified in the study 20180065 protocol. The data will be managed according to the Data Management Plan.

22.3 Handling of Missing and Incomplete Data

Assumptions for missing or partially missing dates for important variables will be made to allow inclusion of appropriate data records in the analyses.

Imputed dates will not be presented in the listings. However, if the partial AE onset date information does not indicate whether the AE started prior to treatment or after the TEAE period, the AE will be classified as TEAEs.

The handling of incomplete and partial dates for adverse events, concomitant medications, death, **subsequent** anti-cancer therapy, **initial ALL diagnosis, prior ALL therapy medication, prior HSCT, remission to prior ALL therapy, and relapse to prior ALL therapy** are described in Appendix F.

The following imputation rules may be considered for the baseline missing (or unknown) value of covariates within **each phenotype and** the treatment arm where it occurred: for continuous covariates, the mean of the observed data will be considered, while for categorical covariate, the missing/unknown values will be combined with the category observed on the most subjects if the missing/unknown rate is **at most** 10% of the observed data, or they will be classified into a separate category if the missing/unknown rate is **more than** 10%. A sensitivity analysis may be conducted to evaluate the primary analysis outcome based on a threshold other than 10% for a separate missing category. If the covariates are also used as subgroup variables, the imputed values of subgroup variables will not be used for subgroup categorization.

22.4 Outliers

Pharmacokinetic (PK) [plasma] concentration data will be evaluated for outliers, and decisions to re-assay individual samples will be made in accordance with standard pharmacokinetic evaluation practice.

22.5 Distributional Characteristics

Where appropriate, the assumptions underlying the proposed statistical methodologies will be assessed. If required, data transformations or alternative non-parametric methods of analyses will be utilized.

22.6 Validation of Statistical Analyses

Programs will be developed, maintained, and the output will be verified in accordance with current risk-based quality control procedures.

Tables, figures, and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analyses consists of Amgen-supported versions of statistical analysis software; for example, the SAS System version 9.4 or later.

23. Statistical Methods of Analysis

23.1 General Considerations

Unless otherwise specified, the data will be presented for the Safety Analysis Set for B-cell and T-cell phenotypes analyzed separately.

Summary statistics will be provided for selected endpoints. For continuous variables, the number of subjects with non-missing data (n), mean, standard deviation, median, minimum, and maximum will be presented. For discrete data, the frequency and percent distribution will be presented. Unless otherwise indicated, percentages will be calculated with respect to the number of subjects in the analysis set. Confidence intervals, when presented, will be constructed at the 2-sided 95% level. For binomial variables, exact distribution methods will be employed. The distribution of time-to-event endpoints will be summarized by Kaplan-Meier method when applicable.

Individual subject listings will be produced for selected endpoints using data recorded on the electronic case report forms (eCRFs) or derived data.

23.2 Detection of Bias

Lack of protocol compliance and the potential for biased statistical analyses will be examined by assessing the incidence of important protocol deviations. The clinical study team will identify and document the criteria for important protocol deviations.

23.3 Subject Accountability

The following subject disposition information will be summarized descriptively:

- Number of enrolled subjects
- Number of subjects **with** screen failures
- Number of treated subjects (Safety population)
    - Number (%) of subjects who discontinued **study** treatment **(including the protocol approved alternatives)** and the primary reason for discontinuation. The percentage is calculated with respect to Safety population.
    - Number (%) of subjects who **completed the study, and who** discontinued study and the primary reason for discontinuation. The percentage is calculated with respect to Safety population.
    - Number (%) of subjects who discontinued each study treatment due to COVID-19 control measures.

Key study dates for the first subject enrollment, last subject enrollment, **last subject last dose of investigational product, last subject end of study** and data cut-off date will be presented.

The number (%) of subjects who were enrolled will be tabulated by region, country, and investigator site, by age group defined in Section 1**8**.2 for safety analysis set.

MRD sample disposition will be summarized for subjects after induction therapy (Induction Safety Analysis Set) and consolidation therapy (Consolidation Safety Analysis Set) separately. The number (%) of subjects with MRD sample at baseline/post-baseline, and the number (%) of subjects without MRD sample at baseline/post-baseline along with the reasons will be presented.

In addition, the number (%) of subjects in each analysis set will be summarized, including the external control arm.

23.4 Important Protocol Deviations

Important Protocol Deviations (IPDs) categories potentially affecting efficacy and safety are defined by the study team before the first subject's initial visit and updated during the IPD reviews throughout the study prior to database lock. During the course of the study these definitions will be used to identify the IPDs that will be summarized at primary and final analysis. Additionally, the major IPDs that will exclude patients from Per Protocol analysis set will be clarified prior to any planned analysis. Eligibility deviations are defined in the protocol. The following information of IPD and protocol deviation will be summarized, where applicable, with respect to the following:

- Number (%) of subjects with IPDs and total number of IPDs will be summarized by category and sub-category
- Number (%) of subjects with IPDs related to COVID-19 control measures will be summarized
- Number (%) of subjects with COVID-19 protocol deviations by protocol deviation category (940 series, 950 series and 960 series of protocol deviations codes)
- Subject listing of IPD in Safety Analysis Set, including COVID-19 related IPDs with the descriptions.
- Subject listing of COVID-19 protocol deviations based on protocol deviations category (940 series, 950 series and 960 series of protocol deviations codes)

23.5 Demographic and Baseline Characteristics

Demographic (age **at diagnosis (years): continuous; with categories of <1, 1 to < 2, 2 to < 10, ≥ 10; age at study entry: continuous (years); with categories of <1 month, 1 month - ≤ 17 years, > 17 years,** sex, ethnicity, race, region), baseline characteristics (weight (kg), height (cm), BSA (m^2), and BMI (kg/m^2) at baseline) and baseline disease characteristics (defined in Section **11**.4.3) will be summarized using descriptive statistics **in Safety Analysis Set**. **The baseline characteristics for subgroups (Section 18.2) and** [redacted]

Additional demographics summary will be presented in side-by-side for phase 1 total, phase 2 total, and total of phase 1 and phase 2 data.

23.6 Efficacy Analyses

Unless otherwise specified, all efficacy endpoints will be analyzed for B-cell and T-cell phenotype independently. The details on endpoint definitions are included in Section 1**6**.1 **and Section 19**. The comparison analyses for CR, CRh (or better), CRp (or better), CRi (or better), EFS, and OS endpoints will use the PAS, while for all the other efficacy endpoints the Safety Analysis Set will be used. DOR will be analyzed only on responders from PAS (i.e., subjects who reached a response of CR, CRp, CRh, or CRi).

Table 23.1. Primary Efficacy Endpoint Summary Table

Endpoint	Primary Summary and Analysis Method	Sensitivity Analysis
CR after induction therapy:	As described in [redacted] The odds ratio (OR) for the probability of CR (PoCR) for CFZ-VXLD versus control will be estimated from a logistic regression model with inverse probability treatment weights (IPTW) for the average treatment effect of the treated (ATT). These weights are derived for each CFZ-VXLD and control subject based on their propensity score (PS) for receiving CFZ-VXLD. Outlier weights will be trimmed via a tuning process. The 95% confidence interval of the OR will be calculated using robust sandwich covariance estimator and the lower bound of the 95% CI will be compared to 1. The CR rate and 95% CI by arm will also be estimated using the same logistic regression model. The CR **rate** at the end of Induction will also be summarized **descriptively** along with the Clopper Pearson 95% CI without any PS adjustment **for Induction Safety Analysis Set and Per Protocol Set**.	• **Treatment response based on investigator assessment** • **Sponsor derived treatment response using central lab data** • Outliers in weights will be trimmed to different threshold values • Add additional potential prognostic covariates in PS model • Rule-out method • Chemotherapy Analysis Set (CAS) • VXLD/VPLD Analysis Set (VAS) • Covariates that remain imbalanced

Endpoint	Primary Summary and Analysis Method	Sensitivity Analysis
		after PS adjustment will be added in the analysis model • Missing baseline covariate imputed with a threshold other than 10% for a separate missing category • Model selection algorithm may be applied to the pre-specified covariates

Table 23.2. Secondary Efficacy Endpoint Summary Table

Endpoint	Primary Summary and Analysis Method	Sensitivity Analysis
CRi or better at the end of induction therapy	Similar to primary endpoint, the rate of CRi or better per arm and OR for CFZ-VXLD versus external control, along with their 95% CI will be estimated from weighted models as for the primary endpoint analysis via IPTW-ATT. The unadjusted rate and 95%CI will **also** be estimated for CFZ-VXLD arm without any PS adjustment **for Induction Safety Analysis Set and Per Protocol Set**.	• **Treatment response based on investigator assessment** • **Sponsor derived treatment response using central lab data** • Outliers in weights will be trimmed to different threshold values • Add additional potential prognostic covariates in PS model • Rule-out method • CAS • VAS • Covariates that remain imbalanced after PS adjustment will be added in the analysis model. • Missing baseline covariate imputed with a threshold other than 10% for a separate missing category • Model selection algorithm may be

		applied to the pre-specified covariates
EFS	EFS will be summarized by Kaplan-Meier method. The distribution of EFS time including median and other quartiles will be summarized descriptively using the KM method. The corresponding 95% CI for the median and other quartiles will be constructed using KM method with log-log transformation. Survival rates at 6, 12, 18, **and** 24 months will be estimated, and the corresponding 95% CIs will be calculated using the method of Kalbfleisch and Prentice (1980). The duration of the follow-up for EFS will be estimated by reverse Kaplan-Meier method (Schemper and Smith 1996). KM curves will also be presented. The hazard ratio (HR) for CFZ-VXLD vs external control arm will be derived using Cox regression. This analysis will be adjusted following a similar approach and weights as for primary endpoint analysis. **The above analyses will also be performed for CFZ-VXLD arm without any PS adjustment for Safety Analysis Set and Per Protocol Set.**	• **Treatment response based on investigator assessment** • Outliers in weights will be trimmed to different threshold values • Add additional potential prognostic covariates in PS model • Rule-out method • CAS • VAS • Same analysis approach as for the primary analysis with control subjects being censored at 2 years since initiation of therapy, only if the follow-up time is different between treatment and control arms. • Covariates that remain imbalanced after PS adjustment will be added in the analysis model. • Missing baseline covariate imputed with a threshold other than 10% for a separate missing category • Model selection algorithm may be applied to the pre-specified covariates

OS	OS will be summarized by Kaplan-Meier method. The distribution of OS time including median and other quartiles will be summarized descriptively using the KM method. The corresponding 95% CI for the median and other quartiles will be constructed using KM method with log-log transformation. Survival rates at 6, 12, 18, **and** 24 months will be estimated, and the corresponding 95% CIs will be calculated using the method of Kalbfleisch and Prentice (1980). The duration of the follow-up for OS will be estimated by reverse Kaplan-Meier method (Schemper and Smith 1996). KM curves will also be presented. The hazard ratio (HR) for CFZ-VXLD vs external control arm will be derived using Cox regression. This analysis will be adjusted following a similar approach and weights as for primary endpoint analysis. **The above analyses will also be performed for CFZ-VXLD arm without any PS adjustment for Safety Analysis Set and Per Protocol Set.**	• Outliers in weights will be trimmed to different threshold values • Add additional potential prognostic covariates in PS model • Rule-out method • CAS • VAS • Same analysis approach as for the primary analysis with control subjects being censored at 2 years since initiation of therapy, only if the follow-up time is different between treatment and control arms • Covariates that remain imbalanced after PS adjustment will be added in the analysis model. • Missing baseline covariate imputed with a threshold other than 10% for a separate missing category • Model selection algorithm may be applied to the pre-specified covariates

DOR of CRi or better	DOR will be summarized by treatment arm on subjects who reach CRi or better by the end of study treatment using the Kaplan-Meier method. The distribution of DOR time including median and other quartiles will be summarized descriptively using the KM method. The corresponding 95% CI for the median and other quartiles will be constructed using KM method with log-log transformation. Survival rates at 6, 12, 18, and 24 months will be estimated, and the corresponding 95% CIs will be calculated using the method of Kalbfleisch and Prentice (1980). **The above analyses will also be performed for CFZ-VXLD arm without any PS adjustment for Safety Analysis Set and Per Protocol Set.**	• NA
MRD **status** using NGS **less than** 10^{-4} after induction therapy in subjects achieving CR	Estimate the proportion of subjects whose MRD level using NGS is **less than** 10^{-4} after induction therapy in subjects achieving CR. Generate Clopper Pearson 95% confidence interval for CFZ-VXLD arm **(Induction Safety Analysis Set)**.	• NA
MRD **status** using NGS **less than** 10^{-3} and **less than** 10^{-4} in subjects achieving CRi or better, after induction and consolidation therapy, **separately**	Estimate the proportion of subjects whose MRD level using NGS **less than** 10^{-3} and **less than** 10^{-4} separately **in subjects achieving CRi or better,** after induction and consolidation therapy, respectively, and generate Clopper Pearson 95% confidence interval in each case for the subjects in CFZ-VXLD arm **(Induction Safety Analysis Set, Consolidation Safety Analysis Set, respectively).**	• NA
Occurrence of a stem cell transplant or CAR-T, without an intervening relapse after protocol specified therapy	Estimate the proportion of subjects that bridge to stem cell transplant or chimeric antigen receptor T cell. Generate Clopper Pearson 95% confidence interval for CFZ-VXLD arm **(Safety Analysis Set).**	• NA
CRi or better after consolidation therapy	Estimate the proportion of subjects who **began consolidation therapy and** achieved CRi or better after the consolidation therapy and its Clopper Pearson 95% CI for CFZ-VXLD arm **(Consolidation Safety Analysis Set).**	• NA

Table 23.3. Exploratory Efficacy Endpoint Summary Table

[redacted]

23.6.1 Analyses of Primary Efficacy Endpoint(s)

To evaluate the treatment effect with respect to the outcome of CR rate among pediatric R/R ALL subjects, an external control arm representing pediatric ALL subjects receiving SoC without **Carfilzomib** will be selected from the observational Study 20180065 according to the study protocol and the definition of each analysis set as specified in Section **20**. For the primary analysis, the 20140106 study will be analyzed as a virtual randomized controlled trial comparing enrolled subjects treated with the CFZ-VXLD regimen versus the external control arm adjusted via propensity score (PS) methods. To control confounding for measurable variables, each subject in the analysis will have a PS estimated via multivariable logistic regression models constructed CCI [redacted]

The odds ratio (OR) for the probability of CR (PoCR) for CFZ-VXLD versus control will be estimated from a logistic regression model with inverse probability treatment weights for the average treatment effect of the treated (IPTW-ATT). These weights are derived for each CFZ-VXLD and control subject based on the PS for receiving CFZ-VXLD. Conditioned on these weights, the distribution of the observed baseline covariates will be balanced

between the 2 arms. If identified, outlier weights will be trimmed to a maximum non-outlying value (i.e., max of 5 or 99th percentile of weight distribution) (Lee, 2011).

The 95% confidence interval (CI) of the OR will be calculated using a robust sandwich covariance estimator and the lower bound of the 95% CI will be compared to 1. The CR rate and 95% CI by arm will also be estimated using the same logistic regression model.

[redacted].

Estimated CR rate and 95% CI by arm along with OR (95% CI) of CR between appropriately weighted CFX-VXLD and external control using PS will be calculated via the weighted logistic regression for the subgroups listed in Section 1**8**.2. The same weights as those included in the primary analysis will be used for subgroup analyses. A forest plot will be produced for the estimated ORs (95% CIs). [redacted] If the balance criteria are not met, then a sensitivity analysis will be considered for estimating the OR (95% CI) using IPTW adjusted models constructed only considering subjects in the subgroup. The re-adjusted summaries would provide ancillary information on how well balance is or is not achieved within each subgroup and may help aid in the interpretation. This sensitivity analysis will be done only if the number of subjects within each category (for the number of subjects in a subgroup <20%, the category is for the combined category) is more than 40% from overall sample size per arm. The CFZ-VXLD subjects reaching CR at the end of Induction will also be summarized along with their Clopper Pearson 95% CI without any PS adjustment.

Several sensitivity analyses will be considered to evaluate the robustness of the results.

Depending on the outlier weights identified during the primary endpoint analysis, several threshold values will be used for trimming and applied to the models used for primary endpoint analysis.

In order to address NCCN recommendation for standard of care therapies in pediatric ALL, a sensitivity analysis will be performed on Chemotherapy Analysis Set (CAS). Similar adjustment as for the primary analysis will be considered.

Despite the attempt to balance covariates using a PS strategy, possible residual bias due to unmeasured confounding may still exist. In order to address this limitation, additional sensitivity analyses will be considered. One will expand the list of covariates to a larger set, [redacted] Another one will apply quantitative bias

analyses to assess the extent of unmeasured confounding that would be required to refute an observed difference in outcome incidence between cohorts. The rule-out method is used **in the quantitative bias analyses** to assess how strong a confounder (or set of confounders) would need to be to fully explain the observed association between an exposure and the outcome. This method has been described previously (Schneeweiss et al 2006), is publicly available at www.drugepi.org and HR variant has been previously applied in the literature (Weintraub et al 2012). An alternative method of evaluating unmeasured confounding involves assessing the strength of the measured confounders by removing each confounder individually from the model to develop a distribution of the point estimate of the hazard ratios to display the strength of the measured confounding. Assuming the unmeasured confounders are similarly distributed, this distribution can be used to inform the potential magnitude and direction of the unmeasured confounders on the validity of the effect estimate. One or more of these methods will be explored, as appropriate. Covariates that remain imbalanced after PS adjustment will be added in the analysis model as a sensitivity analysis. A sensitivity analysis will be performed for missing baseline covariate imputed with a threshold other than 10% for separate missing category. Model selection algorithm may be applied to the pre-specified covariates as a sensitivity analysis, [redacted]

Another sensitivity analysis will be performed for VXLD/VPLD analysis set by estimating the CRR and 95%CI by arm with no PS adjustment.

In addition, historical estimates of CR in patients treated with best available therapy will be provided via a literature search. Descriptive summaries will be made in order to evaluate whether the experimental arm provides benefit over available therapy.

23.6.2 Analyses of Secondary Efficacy Endpoint(s)

The proportion of safety analysis subjects and corresponding Clopper Pearson 95% CI will be estimated for the below secondary endpoints. **The analysis set used for each endpoint is specified in Table 23.2.**

- MRD **status using NGS less than** 10^{-4} after induction therapy in subjects achieving CR
- Occurrence of a stem cell transplant or CAR-T, without an intervening relapse after protocol specified therapy
- **CRi, CRh, CRp, or CR after induction therapy**

- **CRi, CRh, CRp, or CR** after consolidation therapy
- MRD **status using NGS less than** 10^{-3} and **less than** 10^{-4} in subjects achieving CRi **CRi, CRh, CRp, or CR** after induction and consolidation therapy, separately

All secondary endpoints that will be included in the comparative analyses will be included in adjusted models similarly as the primary endpoint.

The rate of CR, CRp, CRh and CRi at the end of induction **and at the end of consolidation** per arm and the OR for CFZ-VXLD versus control will be estimated from a logistic regression model with IPTW-ATT adjustment along with their 95% CIs. This analysis will be performed following a similar approach and sensitivity analyses as for primary endpoint analysis.

EFS, OS, and DOR will be summarized by Kaplan-Meier method. The distribution of EFS/OS/DOR time including median and other quartiles will be summarized descriptively using the KM method (Klein and Moeschberger, 1997). The corresponding 95% CI for the median and other quartiles will be constructed using KM method with log-log transformation. Survival rates at 6, 12, 18, and 24 months will be estimated for EFS/OS, and the corresponding 95% CIs will be calculated using the method of Kalbfleisch and Prentice (1980). The duration of the follow-up will be estimated by reverse Kaplan-Meier method (Schemper and Smith 1996). KM curves will also be presented for OS and EFS. The hazard ratio (HR) for CFZ-VXLD vs external control arm will be derived for OS/ EFS using Cox regression. This analysis will be adjusted following a similar approach and weights as for primary endpoint analysis. If the sample size is small and/ or censoring is high, then the KM method may not provide reliable estimates **for DOR**, and in this case, only descriptive statistics or listings will be provided.

Similar sensitivity analyses as for primary endpoint analysis will be considered for EFS/OS, as described in Table **23**.2. Additionally, if the follow-up time from control arm is different than experimental arm, then a sensitivity analysis similar to the primary analysis will be performed with control subjects being censored at 2 years since initiation of therapy.

The proportion of subjects and corresponding Clopper Pearson 95% CI will be estimated for MRD status in subjects achieving CR and CRi or better (CRi, CRp, CRh, CR) as described in Table 23.2.

23.6.3 Analyses of Exploratory Efficacy Endpoint(s)

23.7 Safety Analyses

23.7.1 Analyses of Safety Endpoint(s)

The safety analysis will be based on the safety analysis set analyzed by B-cell, T-cell separately and pooled. The summaries will be presented in side-by-side displays for B-cell/ T-cell/ Total (i.e., pooled sample). **The safety analysis will be also performed** for Induction and Consolidation **phase** separately.

The analysis for induction treatment period will be based on the Induction Safety Analysis Set (Section 20.2). TEAEs that started within the first dose of induction treatment (inclusive) to the earlier of 30 days after the last dose date of any study treatment, or the first dose of consolidation treatment (not inclusive) will be included in this analysis.

The analysis for consolidation treatment period will be based on the Consolidation Safety Analysis Set (Section 20.2). TEAEs that started within the first dose of consolidation treatment (inclusive) to the earlier of the end of study (inclusive), or 30 days after the last dose date of any study treatment will be included in this analysis.

The analysis will include the descriptive summary statistics for Adverse events, Laboratory measures, Vital Signs, Physical measurements, Karnofsky or Lansky Performance Status, ECHO and ECG by age group defined in Section 1**8**.2. Details are described in respective section below.

23.7.2 Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) version **25.**1 or later will be used to code all events categorized as adverse events to a system organ class and a preferred

term. **The events of interest (EOI) search strategies will be based on the standardized MedDRA query (SMQ) and/or Amgen MedDRA query (AMQ). Incomplete AE start dates will be imputed according to the specifications described in Appendix F1.**

The subject incidence of TEAEs will be summarized for all treatment-emergent adverse events (TEAEs), TEAEs of grade 3 or higher, treatment related **TE**AEs, serious **TEAEs** (SAEs), **treatment related SAEs, TE**AEs leading to discontinuation of study treatment, fatal **TE**AEs and treatment-emergent EOI, tabulated by system organ class (SOC) in alphabetical order and preferred term (PT) in descending order of frequency and by PT in descending order of frequency without regard to SOC. The maximum NCI – CTCAE 4.03 toxicity grade and strongest causal relationship to study treatment will **be** evaluated, and the TEAEs will be also summarized by severity and relationship to the study drugs.

If a subject experiences repeated episode of the same **TE**AE, the subject will be counted only once within each SOC and similarly counted once within each PT and the event with the highest severity grade will be used for purposes of incidence tabulations.

A summary of the number of deaths and the cause of death, classified by deaths within 30 days of last dose of study drug and deaths more than 30 days after the last dose of study drug, will be provided. A listing of deaths will be provided.

23.7.3 Laboratory Test Results

Laboratory test results will be graded for severity using the NCI-CTCAE version 4.03 and will be summarized using descriptive statistics for baseline values and changes from baseline values by assessment visit (per Table 14, Table 16, Table 18 and Table 20 from protocol), and a summary of subject incidence of grade 3 and 4 laboratory abnormalities for all laboratory parameters of interest.

Shifts in laboratory toxicity grades to outside the normal range will be evaluated for laboratory parameters by assessing the maximum increase and/or decrease observed during the course of study treatment relative to the baseline toxicity grade.

For the summary of changes from baseline values, subjects without a baseline and/or post-baseline value will be excluded. Laboratory results from samples taken $>$ 30 days after the last administration of protocol therapy will be excluded from all laboratory summaries.

The laboratory analytes of interest are:

- CBC with differential (i.e., Total Neutrophils, Eosinophil, Basophils, Lymphocytes, Monocytes, Lymphoblasts) and platelet count
- Electrolytes: Sodium, Potassium, Chloride, Bicarbonate, Calcium, Magnesium, Phosphorus
- Serum glucose
- Renal function: blood urea nitrogen (BUN), Creatinine
- Hepatic function: AST, ALT, total and direct bilirubin, ALP, albumin
- Pancreatic function: amylase, lipase
- Uric acid

The subject incidence and percentage of potential Hy's Law cases will also be summarized.

23.7.4 Vital Signs

Actual value and change from baseline for vital sign results including systolic/diastolic blood pressure, **heart rate**, respiratory rate, and temperature will be summarized using descriptive statistics for baseline values and changes from baseline by assessment visit (per Table 14, Table 16, Table 18 and Table 20 from protocol).

For the summary of changes from baseline, subjects without a baseline and/or post-baseline value will be excluded. Vital sign results taken > 30 days after the last administration of protocol therapy will be excluded from all vital sign summaries.

23.7.5 Physical Measurements

Actual value and change from baseline for the physical measurements including height(cm), weight(kg), and BSA(m^2) will be summarized using descriptive statistics by assessment visit (per Table 14, Table 16, Table 18 and Table 20 from protocol). For the summary of changes from baseline, subjects without a baseline and/or post-baseline value will be excluded. Physical measurements results taken > 30 days after the last administration of protocol therapy will be excluded from all physical measurements summaries.

23.7.6 Karnofsky or Lansky Performance Status

Shifts in scores for Karnofsky (age >16 years), Lansky (<16 years) performance status scores between screening and safety follow up will be tabulated. Performance status will be assessed by the Karnofsky and Lansky scales per Section **24**.2.1.7 of the protocol.

23.7.7 Electrocardiogram

The electrocardiogram (ECG) measurements from this clinical study were performed as per standard of care for routine safety monitoring, **rather than for the purposes of assessment of potential QTc**. ECG data **may** be presented in listings.

23.7.8 Echocardiogram

Echocardiogram, including assessments of systolic and diastolic left ventricular function and right ventricular function, will be summarized at baseline using descriptive statistics based on the Safety Analysis Set. **Additionally, the unscheduled ECHO assessments (as clinically indicated, such as congestive heart failure) will be summarized using descriptive statistics, based on the number of adverse events that trigger the unscheduled ECHO assessments.**

23.7.9 Antibody Formation

Not Applicable

23.7.10 Exposure to Investigational Product

Exposure will be summarized separately for induction therapy (per Induction Safety Analysis Set) and consolidation therapy (per Consolidation Safety Analysis Set). Descriptive statistics in terms of treatment duration (days), number of actual treatment days, total dose received, number of doses administered, average dose per administration, actual dose intensity, and relative dose intensity will be produced to describe the exposure to carfilzomib. Number (%) of subjects with carfilzomib dose modifications (missed doses, dose delays, dose reductions, and dose interruptions, if applicable) will be summarized. If the reason for dose modification is due to COVID-19 control measures which is reported on the CRF, then the number (%) with COVID-19 control measures will also be presented.

Descriptive statistics will be produced to describe the exposure to investigational product.

23.7.11 Exposure to Non-investigational Product

All the analyses planned for IP will be carried over for each drug in any chemotherapy combination **(VXLD and modified BFM therapy including protocol approved alternative agents) except that the total dose received, average dose per administration, actual dose intensity and relative dose intensity will not include the alternative agents**. For **IT chemotherapy** drugs (methotrexate, cytarabine, **and hydrocortisone**), **only (1) number of doses administered; and (2) ratio of number of actual doses to number of the protocol specified doses** will be summarized **and presented** for CNS positive and negative separately. **Exposure to non-IP will also be**

summarized separately for induction therapy (per Induction Safety Analysis Set) and consolidation therapy (per Consolidation Safety Analysis Set).

In addition, the number and percentage of subjects who received V (vincristine), X (dexamethasone), L (PEG-asparaginase) and D (daunorubicin), or an approved alternate agent, in induction chemotherapy will be summarized for each agent.

23.7.12 Exposure to Concomitant Medication

The number and proportion of subjects receiving therapies of interest will be summarized by preferred term or category as coded by the World Health Organization Drug (WHO DRUG) dictionary.

23.8 Other Analyses

23.8.1 Analyses of Pharmacokinetic or Pharmacokinetic/ Endpoints

Pharmacokinetic analyses will be conducted as was done for the phase 1b part of the study (please refer to Section 11.7.1).

23.9 Changes from Protocol-specified Analyses

There are no changes to the protocol-specified analyses.

24. Literature Citations / References

Clopper CJ and Pearson ES (1934) The use of confidence or fiducial limits illustrated in the case of the binomial, Biometrika. 26(4):404-413.

Kalbfleisch, J. and Prentice, R. (1980) The Statistical Analysis of Failure Time Data. John Wiley, New York.

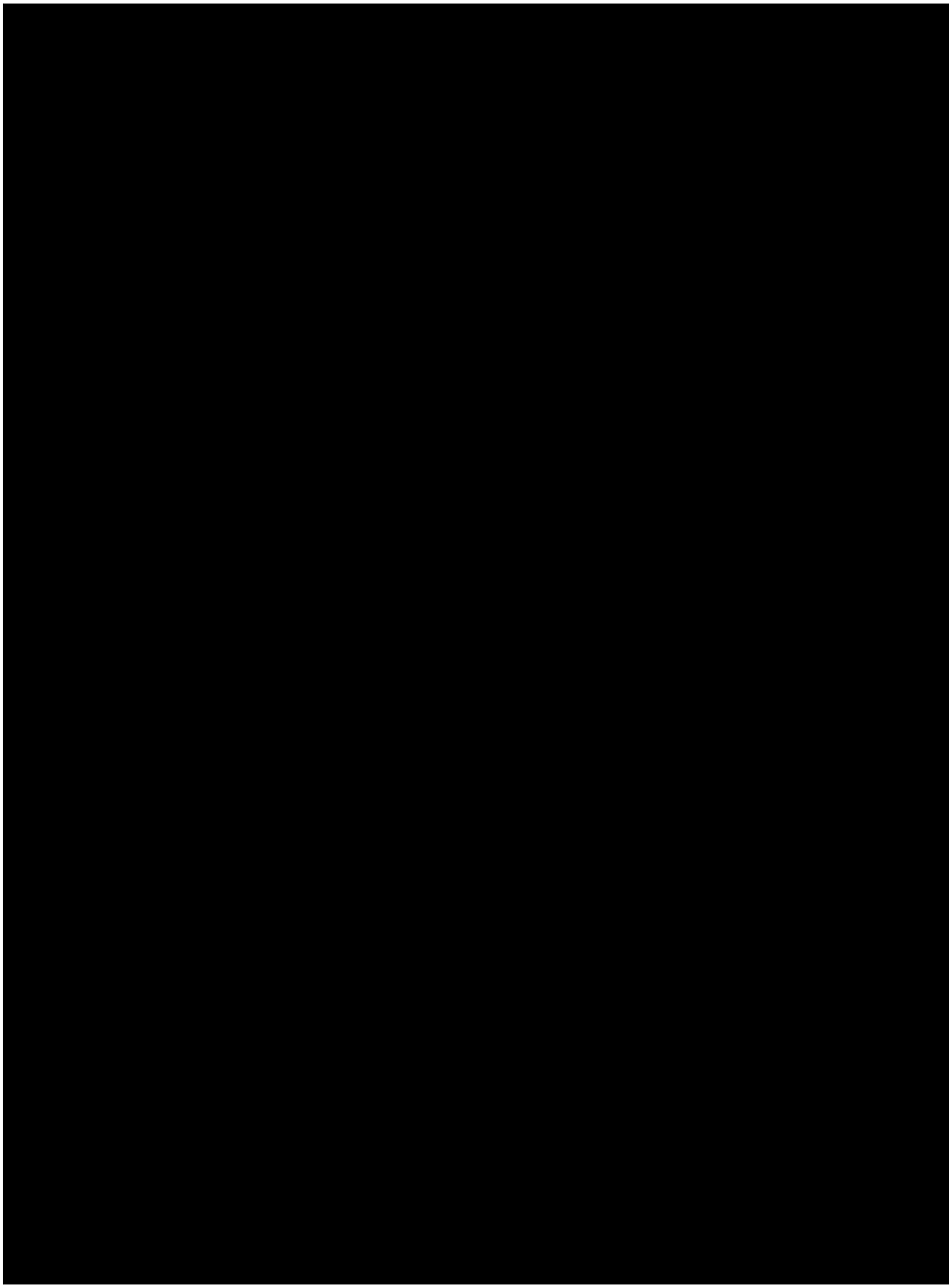
Klein and Moeschberger (1997) Survival Analysis Techniques for Censored and truncated data, Springer.

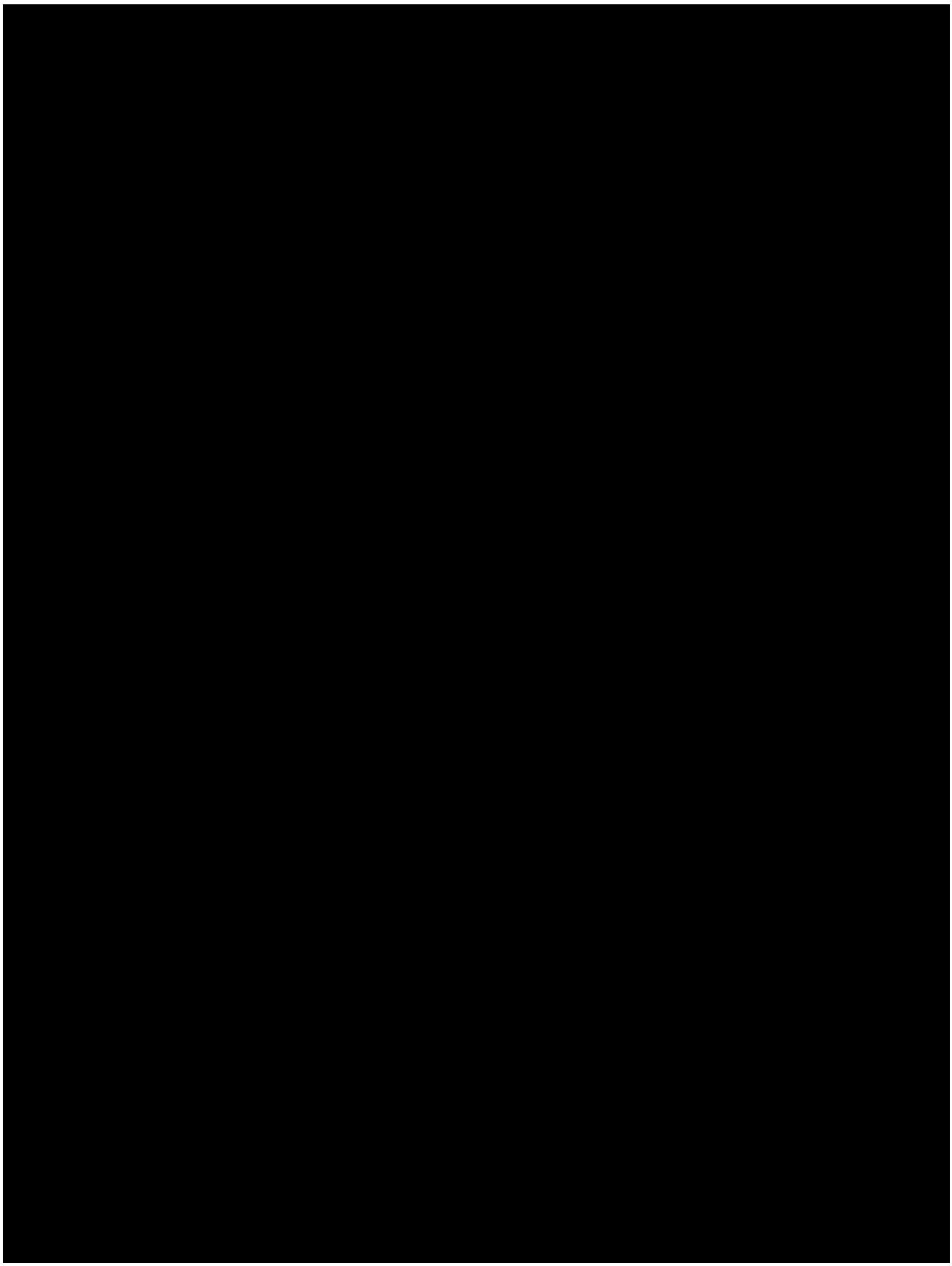
Schemper and Smith (1996) A Note on Quantifying Follow-up in Studies of Failure Time. Controlled Clinical Trials 17:343 – 346

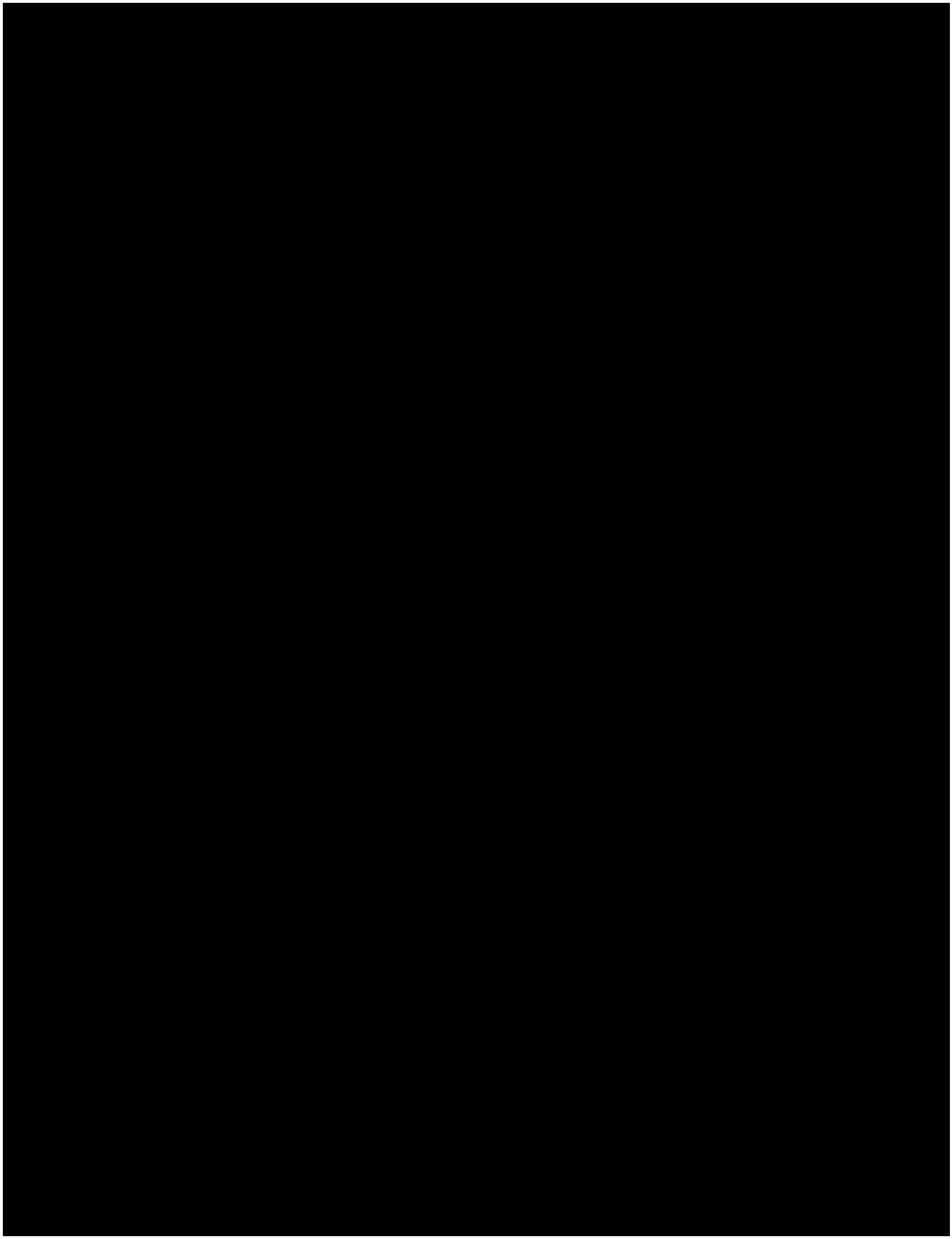
Schneeweiss S. Sensitivity analyses and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. Pharmacoepi Drug Safety 2006; 15:291–303

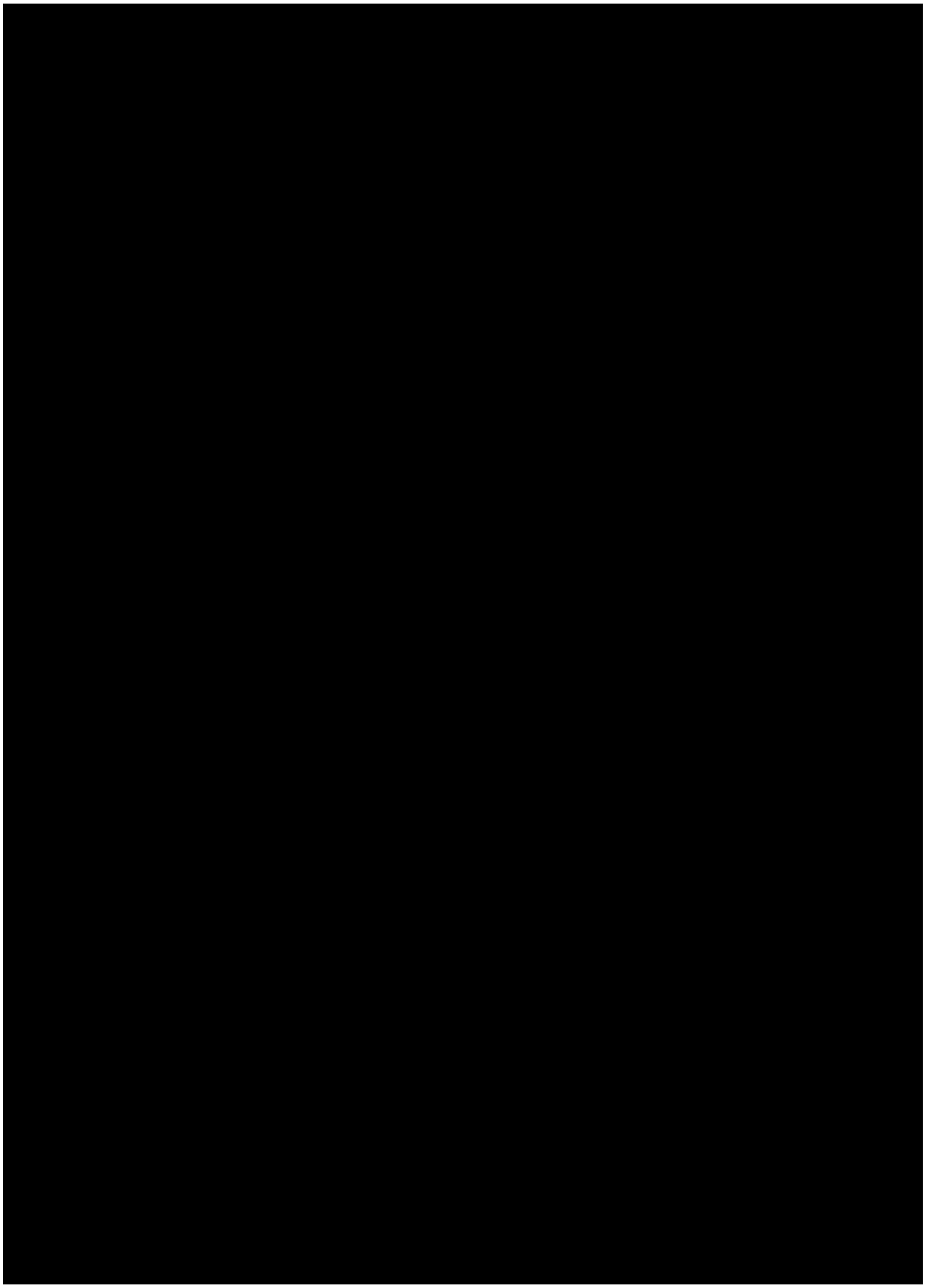
Weintraub WS et al Comparative Effectiveness of Revascularization Strategies. N Engl J Med. 2012 Apr 19; 366(16): 1467–1476.

[redacted]



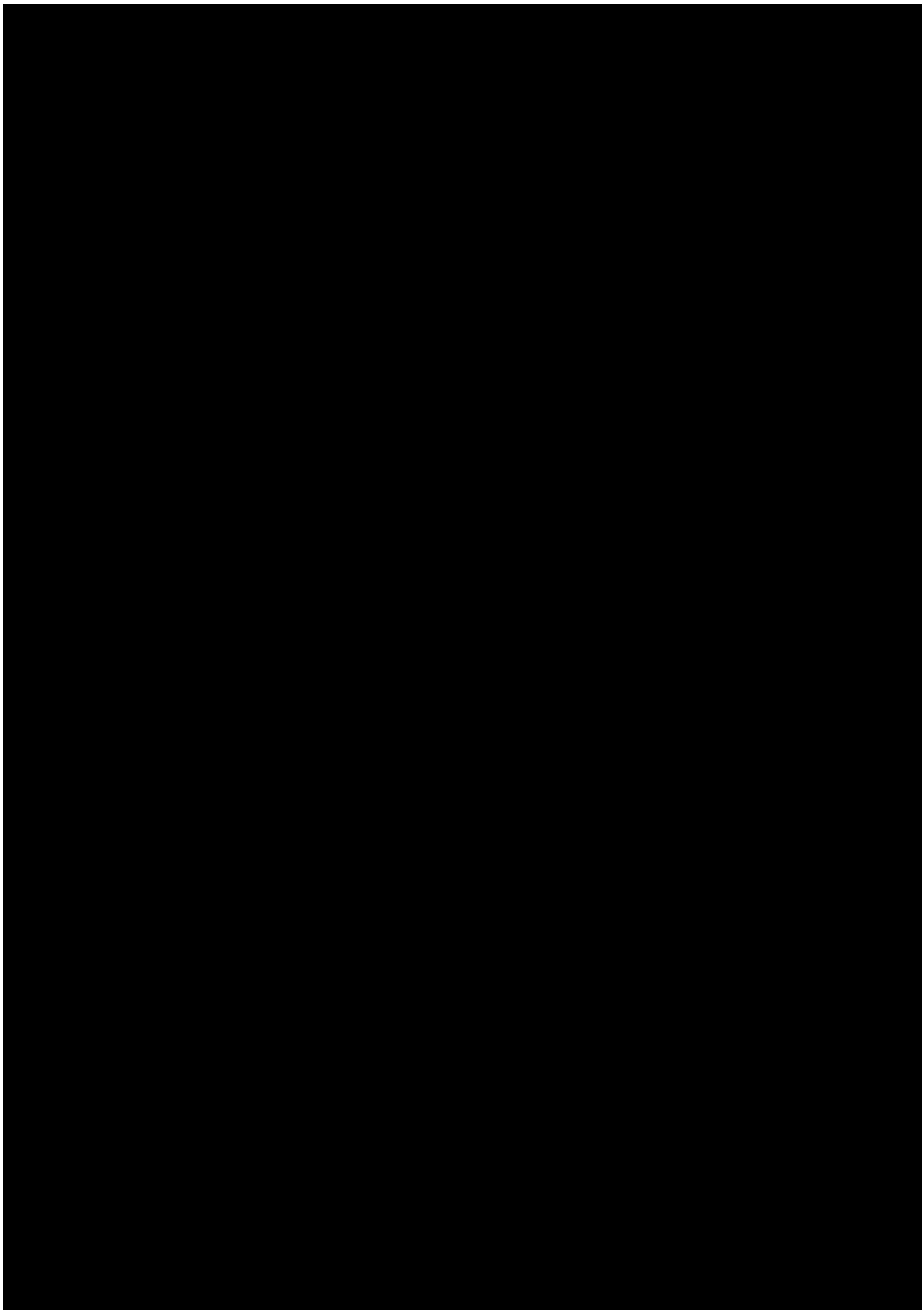






[redacted]

[redacted]



[redacted]

Appendix F. Handling of Incomplete Dates and Missing Dates

F1. Imputation Rules for Adverse Events and Concomitant Medications (other than the new anti-leukemia therapy) dates

The following data will be imputed using the following algorithm:

Adverse Events

Concomitant Medications (other than anti-cancer therapy)

Table F-1. Imputation Rules for Partial or Missing Start Dates

Start Date		Stop Date: Complete: *yyyymmdd*		Stop Date: Partial: *yyyymm*		Stop Date: Partial: *yyyy*		
		$< 1^{st}$ dose	$\geq 1^{st}$ dose	$< 1^{st}$ dose *yyyymm*	$\geq 1^{st}$ dose *yyyymm*	$< 1^{st}$ dose *yyyy*	$\geq 1^{st}$ dose *yyyy*	missing
Partial: *yyyymm*	$= 1^{st}$ dose *yyyymm*	2	1	**n/a**	1	n/a	1	1
	$\neq 1^{st}$ dose *yyyymm*		2		2	2	2	2
Partial: *yyyy*	$= 1^{st}$ dose *yyyy*	3	1	3	1	n/a	1	1
	$\neq 1^{st}$ dose *yyyy*		3		3	3	3	3
Missing		4	1	4	1	4	1	1

1 = Impute the date of first dose
2 = Impute the first day of the month
3 = Impute January 1 of the year
4 = Impute January 1 of the stop year

Note: If the start date imputation leads to a start date that is after the stop date, then do not impute the start date.

Imputation rules for partial or missing stop dates:

For partial stop date mmyyyy, impute the last day of the month.

For partial stop date yyyy, impute December 31 of the year.

For completely missing stop date, do not impute.

If the stop date imputation leads to a stop date that is after the death date, then impute the stop date as the death date.

If the stop date imputation leads to a stop date that is before the start date, then there is a data error and do not impute the stop date. (i.e., set the stop date as missing).

F2. Imputation of the Subsequent Anti-cancer Therapy Start Date

The **subsequent** anti-cancer therapy **start** date will be imputed using the following algorithm. If the start day of **subsequent** anti-cancer therapy is missing and month and year are not the same as last dosing date of study treatment, it will be assumed to be the first day of the month. If the start day of **subsequent** anti-cancer therapy is missing and month and year are same as last dosing date of study treatment, the start date will be assumed as last dosing date of study treatment. In other situations, do not impute.

F3. Imputation of Partial/Missing Death Dates

The following rules will be applied to impute partial or missing death dates.

- If death year and month are available but day is missing:
  - If YYYYMM for the date last known to be alive equals YYYYMM for death date, set death date to the day after the date last known to be alive.
  - If YYYYMM for the date last known to be alive is less than the YYYYMM for death date, set death date to the first day of the death month.

    (Below two situations will be evaluated at time of analysis by using a more conservative way)
  - [If YYYYMM for the date last known to be alive is greater than YYYYMM for death date, assume date last known to be alive is in error, set death date to the first day of the death month.

    or
  - If YYYYMM for the date last known to be alive is greater than YYYYMM for death date, data error, do not impute and censor subject survival time.]
  - If month and day are missing and year of death is known:
    - If YYYY for the date last known to be alive equals the YYYY for death date, set death date to the day after last known to be alive date.
    - If YYYY for the date last known to be alive is less than the YYYY for death date, set death date to the first day of the death year.
    - If YYYY for the date last known to be alive is greater than YYYY for death date, data error, do not impute and censor the subject survival time.
  - If a death date is totally missing, do not impute and censor the subject survival time.

F4. Imputation for Date of Initial ALL Diagnosis

If the day is missing but month and year are available, then impute the day to 1st of the month. If month and/or year is missing or the date is completely missing, then do not impute.

F5. Imputation for Start Date and End Date of Prior ALL Therapy Medication, Date of Prior Hematopoietic Stem Cell Transplant, Date of Remission (CRi or Better) to Prior ALL Therapy, and Date of Relapse to Prior ALL Therapy

If the day is missing but month and year are available, then impute the day to 15th of the month. If month and/or year is missing or the date is completely missing, then do not impute.

[redacted]