TITLE PAGE

STATISTICAL ANALYSIS PLAN

Pooled Analysis of CAEL101-301 and CAEL101-302

Version Number: 2.0

Protocol Titles: A Phase 3, Double-Blind, Multicenter Study to Evaluate the Efficacy and Safety of CAEL-101 and Plasma Cell Dyscrasia Treatment Versus Placebo and Plasma Cell Dyscrasia Treatment in Plasma Cell Dyscrasia Treatment-Naïve Patients with Mayo Stage IIIb AL Amyloidosis

A Phase 3, Double-Blind, Multicenter Study to Evaluate the Efficacy and Safety of CAEL-101 and Plasma Cell Dyscrasia Treatment Versus Placebo and Plasma Cell Dyscrasia Treatment in Plasma Cell Dyscrasia Treatment-Naïve Patients with Mayo Stage IIIa AL Amyloidosis

Protocol Numbers: CAEL101-301 and CAEL101-302

Protocol Amendment Numbers: 6.0 (CAEL101- 301) and 7.0 (CAEL101-302)

Compound: Anselamimab (also referred to as CAEL-101)

Short Title: Cardiac Amyloid Reaching for Extended Survival

Sponsor Name: Alexion Pharmaceuticals, Inc

Legal Registered Address:

121 Seaport Boulevard

Boston MA 02210

USA

Author: PPD

Version Date: 15 May 2025

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

VERSION HISTORY

This statistical analysis plan (SAP) is based on Protocol CAEL101-302 Version 7.0 and Protocol CAEL101-301 Version 6.0.

The SAP has also been updated to version 2.0, reflecting amendments from the previous version 1.0.

SAP Version	Version Date	Change	Rationale
1.0	07 Nov 2024	Not applicable	Original version
2.0	15 May 2025	**Changes following FDA's feedback**	
		Added Text in Section 3	added the power and the corresponding RRR of the pooled analysis under Hazard Ratio=CCI.
		Added text in Section 5.4.4	added a sensitivity analysis using ANCOVA method based on imputed data.
		Added text in Section 6.2.1	Clarified the use of data with respect to the intake of prohibited medication.
		Changes related to secondary endpoint analysis	
		Updated text in Section 5.1.4 and 5.4	Updated the sequential order of the key secondary endpoints. The order of testing for key secondary endpoints was revised based on the statistical power for each parameter based on ranked ANCOVA as well as clinical importance of that endpoint.
		Updated text in Section 5.4.1.3	The summary measure of change from baseline (CFB) to Week 50 has been updated to use the Hodges–Lehmann estimate of the median difference. The treatment effect will be estimated using the Hodges–Lehmann method.
		Added text in Section 5.4.2.1	KM estimates (95% CI) and Relative Risk Reduction (RRR) at Week 50 will be presented.
		Updated text in Section 5.4.2.3	Removed LS means ±SE and the difference in LS means between the 2 intervention groups with 95%CI from Ranked ANCOVA results. p-value for statistical significance will be calculated based on Ranked ANCOVA and treatment effect will be estimated using the Hodges Lehman method.
		Removed text in Section 5.4.4	The sensitivity analysis of secondary endpoints (CFB to Week 50) using a Mixed Model for Repeated Measures (MMRM) approach under the assumption of missing at random (MAR) has been removed. Given the anticipated substantial amount of missing data due to death, the MAR assumption is deemed inappropriate. Instead, a new sensitivity analysis will be conducted using an ANCOVA model based on imputed data under 'Missing Not at Random' assumption, to better account for missingness not at random. This is in addition to the existing sensitivity analysis of repeating the main analysis based on Ranked ANCOVA on the PPS.
		Changes to analysis outside of primary and secondary endpoints	
		Removed text in Section 5.5.2.1	The analysis of exploratory endpoints (CFB over time) using a Mixed Model for Repeated Measures (MMRM) approach under the assumption of MAR has been removed. Given the anticipated substantial amount of missing data due to death, the MAR assumption is deemed inappropriate. As a result, these exploratory endpoints will instead be summarized using descriptive statistics.

		Added Text in Section 5.5.1.2	Added Text to clarify time to graded cardiac response will be analyzed by time to CarCR, time to CarVGPR or better and time to CarPR or better
		Added Text in Section 5.5.1.3	Added Text to clarify time to graded renal response will be analyzed by time to renCR, time to renVGPR or better and time to renPR or better
		Added Text in Section 5.5.1.4	Added Text to clarify time to graded hepatic response will be analyzed by time to hepCR, time to hepVGPR or better and time to hepPR or better
		Added Section 5.7.1	Added Ranked ANCOVA will be repeated for CFB to Week 76 for KCCQ-OS, GLS%, 6MWT, SF-36v2 PCS and CFB to Week 74 for NT-proBNP to align with the duration of the PETP (LPI+18M) Added text to repeat the sensitivity analysis using ANCOVA method based on imputed data for CFB to Week 76 for KCCQ-OS, GLS%, 6MWT, SF-36v2 PCS and CFB to Week 74 for NT-proBNP to align with the duration of the PETP (LPI+18M)
		Updated text in Section 5.7.2.1	Added analysis for deep hematologic response since immunofixation results may potentially be confounded by concomitant administration of monoclonal antibodies of IgG kappa isotype, such as daratumumab or Anselamimab.
		Added Section 5.7.4	Added additional analysis of time to composite of ACM or First CVH to address additional clinical interests.
		Added Section 5.7.5	Added additional analysis of time to first CVH to address additional clinical interests.
		Removed text in Section 5.6.2.1	Removed the overall AE summary by treatment, post-treatment period. The PETP period (LPI +18 months) encompasses safety data collected up to 140 days post-final dose or until the End of Treatment, whichever comes first. As there is no significant post-treatment period to evaluate, post-treatment outputs are excluded from the Statistical Analysis Plan (SAP).
		Removed text in Section 5.6.2.5	Removed the summary of time to onset of AE with Grade ≥3 by SOC and PT as it's not a standard requirement and may be generated post-hoc as needed.
		Removed text in Section 5.6.2.6	Removed the summary of HLT of Infusion site reactions and administration site reactions NEC since events within the HLT of "Injection site reactions" are not representative of the important potential risk of "Serious infusion reactions" and are otherwise included in the SMQ tables for "Hypersensitivity" and "Anaphylaxis"
		Clarification and editorial changes to facilitate review	
		Updated text in Section 1.1	Clarified that PK concentrations by visit will be used to describe the PK profile

		Replaced "major protocol deviation" with "important protocol deviation" throughout the document	To make the terminology of protocol deviation severity category consistent in this document
		Removed text related to "Asian subgroup" in Section 5.1.5	This subgroup has already been covered in Section 5.7.7 Subgroup Analyses
		Updated text in Section 5.1.6	Updated to provide more details on what analyses will be included to support registration in Japan
		Added definition of study duration of PETP in Section 5.2	Added text to clarify how to derive study duration of PETP
		Updated text in Section 5.3.3 Updated Table 2 in Section 5.3.3	Updated text to clarify rules to combine the strata if the sample size is extremely low per study intervention group in any of the strata. Added clarifying text for F-S method to account for ties based on censored versus deceased participants in primary endpoint analysis
		Updated text in Section 5.4	Added text to clarify KCCQ-OS is calculated as the mean of the physical limitation score, total symptom score, QoL score, and social limitation score.
		Updated text in Section 5.4.2.2	Replaced "incidence rate ratio" with "incidence risk ratio" since it is the same estimate, but the latter name is generally better known. Clarified frequency of CVHs per year and the corresponding 95% CI will be reported.
		Removed text in Section 5.5.2.1	Removed the plots of LS means over time by treatment. This information will be presented in the tabular format.
		Added text in Section 5.5.2.2	Clarified renal stability will be analyzed in addition to other renal categories.
		Added text in Section 5.5.2.3	Clarified what strata will be used in the F-S method and win ratio calculation for CCI [redacted]. CCI [redacted]
		Added text in Section 5.5.4	Added text to clarify the immunogenicity analyses will be generated on pooled data as well as by study
		Updated text in Section 5.6	Updated text to clarify weight will be collected with other vital signs

		Updated section title of Section 5.6.2.6 Added text in Section 5.6.2.6	Replaced title “Hypersensitivity and Anaphylactic Reactions by SMQ Terms” with “Infusion-Related Reaction”. Added text to clarify that infusion reactions will be assessed using the Standardized MedDRA Query (SMQ) for hypersensitivity and anaphylactic reactions.
		Added text in Section 5.6.3.2	Clarified local laboratory results will be used for a specific visit if central laboratory results are not available
		Added text in 5.6.3.3	Added text to clarify what will be included in the Summary of liver enzymes.
		Removed and Added text in Section 5.7.2.2	Clarified KM plots for time to ACM will be done by study intervention and daratumumab usage (yes or no)
		Added text in Section 5.7.6	Added more details on the analyses on MRI data from Study CAEL101-302
		Added text and tables in Section 6.1.1	Added text to clarify the analysis windowing rules for EOT visits
		Added text in Section 6.2.1 Removed text in Section 6.2.1	Added “or incorrect treatment at any visit” to clarify this scenario could be a possible reason for important protocol deviation. Removed “Incorrect dose at 1 administration: the actual study drug dose differs from the planned study drug dose by more than ± 10%” from the important PD list. While the overall, cumulative IP dose administered of +/- >20% of the cumulative planned dose is important, an incorrect dose administered at one visit that is more than +/- 10% is considered major/not important because an incorrect dose at an isolated visit would not significantly impact the completeness, accuracy, and/or reliability of key study data nor would significantly affect the subject’s rights, safety, or well-being. Small variances in dose will have no clinical impact.
		Added text in Section 6.2.2	Added text to clarify demographic and baseline characteristics and disease characteristics analyses will be done by individual studies
		Throughout document	Minor administrative changes, editorial changes for clarity, minor corrections for consistency and updates in list of references.

APPROVAL SIGNATURES

This document has been electronically signed in Alexion's document management system. Please refer to the last page for signature details.

PPD

Alexion Pharmaceuticals, Inc

Date: DD Mmm YYYY

PPD

Alexion Pharmaceuticals, Inc

Date: DD Mmm YYYY

PPD

Alexion Pharmaceuticals, Inc

Date: DD Mmm YYYY

LIST OF ABBREVIATIONS

The following abbreviation and acronyms are used in this SAP.

Abbreviation	Definition
%CV	percentage of the coefficient of variation
6MWT	6-Minute Walk Test
AAS	ADA Analysis Set
ACM	all-cause mortality
ADA	antidrug antibody
AE	adverse event
AL	amyloid light chain
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BP	blood pressure
BSA	body surface area
CarCR	cardiac complete response
CarNR	cardiac no response
CarPR	cardiac partial response
CarVGPR	cardiac very good partial response
CEAC	Clinical Event Adjudication Committee
CFB	change from baseline
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CR	complete response
CRP	C-reactive protein
CTCAE	Common Terminology Criteria for Adverse Events
CTD	common technical document
cTnT	cardiac troponin T
CVH	cardiovascular hospitalization
dFLC	involved/uninvolved free light-chain difference
ECG	electrocardiogram
eGFR	estimated glomerular filtration rate
EOS	End of Study
EQ-5D-5L	European Quality of Life Health 5-Dimension 5-Level questionnaire
F-S	Finkelstein-Schoenfeld
FLC	free light chain
GGT	gamma-glutamyl transferase
GH	general health

Abbreviation	Definition
GLS%	global longitudinal strain percentage
hepCR	hepatic complete response
hepNR	hepatic no response
hepPR	hepatic partial response
hepVGPR	hepatic very good partial response
HLT	High Level Term
HR	heart rate
ICE	intercurrent event
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
IDMC	Independent Data Monitoring Committee
IQR	interquartile range
ITT	Intent-to-Treat (Analysis Set)
KCCQ	Kansas City Cardiomyopathy Questionnaire
KCCQ-OS	Kansas City Cardiomyopathy Questionnaire-Overall Sore
KM	Kaplan-Meier
LPI+18M	last participant randomized plus 18 months
LS	least-squares
LVAD	left ventricular assist device
MCID	minimum clinically important difference
MCS	mental component summary
MedDRA	Medical Dictionary for Regulatory Activities
MRI	magnetic resonance imaging
NAb	neutralizing antibody
NCI	National Cancer Institute
NEC	not elsewhere classified
NR	no response
NT-proBNP	N-terminal pro-B-type natriuretic peptide
NYHA	New York Heart Association
PCD	plasma cell dyscrasia
PCS	physical component score
PD	pharmacodynamic(s)
PETP	Primary Evaluation Treatment Period
PF	physical functioning
PK	pharmacokinetic(s)
PKAS	Pharmacokinetic Analysis Set
PP	Per-Protocol (Analysis Set)
PR	partial response
PT	Preferred Term
QoL	quality of life
QTcF	QT interval corrected for heart rate by Fridericia's formula

Abbreviation	Definition
renCR	renal complete response
renNR	renal no response
renPR	renal partial response
renVGPR	renal very good partial response
RMST	Restricted Mean Survival Time
SAE	serious adverse event
SAP	statistical analysis plan
SF-36v2	Short Form 36 Health Survey Version 2
SMQs	Standardised MedDRA Queries
SoA	Schedule of Assessments
SOC	System Organ Class
SS	Safety Analysis Set
TEAE	treatment-emergent adverse event
TESAE	treatment-emergent serious adverse event
ULN	upper limit of normal
VAS	visual analog score
WR	win ratio

1. INTRODUCTION

This SAP describes the statistical methods for analyzing the pooled data from the PETP of two Phase 3 studies:

- Protocol CAEL101-301, "A Phase 3, Double-Blind, Multicenter Study to Evaluate the Efficacy and Safety of CAEL-101 and Plasma Cell Dyscrasia Treatment Versus Placebo and Plasma Cell Dyscrasia Treatment in Plasma Cell Dyscrasia Treatment-Naïve Patients with Mayo Stage IIIb AL Amyloidosis."
- Protocol CAEL101-302, "A Phase 3, Double-Blind, Multicenter Study to Evaluate the Efficacy and Safety of CAEL-101 and Plasma Cell Dyscrasia Treatment Versus Placebo and Plasma Cell Dyscrasia Treatment in Plasma Cell Dyscrasia Treatment-Naïve Patients with Mayo Stage IIIa AL Amyloidosis."

The primary analysis will be based on the data pooled from the 2 studies, CAEL101-301 and CAEL101-302, stratified by study. Analysis of the individual studies will also be performed for all primary and secondary endpoints.

Standard data presentation instructions, and table, figure, and listing specifications are contained in the data presentation plan as a separate document.

1.1. Objectives and Estimand and/or Endpoints

The objectives and estimand and/or endpoints for the PETP are as follows:

Objectives	**Estimand and/or Endpoints**
Primary	
To determine if CAEL-101 and treatment for PCD improve overall survival and/or decrease the frequency of CVH in Mayo stage IIIa or IIIb AL amyloidosis patients who are treatment naïve compared to treatment for PCD and placebo	A hierarchical combination of time to ACM and the frequency of CVH
To evaluate the safety and tolerability of CAEL-101 in combination with treatment for PCD compared to treatment for PCD and placebo	• Incidence of TEAEs and TESAEs • CFBs in vital signs, weight, clinical laboratory tests, and 12-lead ECGs
Secondary[a]	
• To assess quality of life (QoL) as measured by the KCCQ-OS	• CFB to Week 50 in the KCCQ-OS
• To determine if CAEL-101 and treatment for PCD improve survival in Mayo stage IIIa or IIIb AL amyloidosis patients who are treatment naïve compared to treatment for PCD and placebo	• Time to ACM
• To assess functional improvement as measured by the distance walked in the 6MWT	• CFB to Week 50 in the 6MWT
• To assess cardiac improvement as measured by GLS%	• CFBs to Week 50 in the GLS%
• To determine if CAEL-101 and treatment for PCD decrease CVHs in Mayo stage IIIa	• Frequency of CVH

Objectives	Estimand and/or Endpoints
or IIIb AL amyloidosis patients who are treatment naïve compared to treatment for PCD and placebo	
• To assess QoL as measured by the SF-36v2® PCS	• CFBs to Week 50 in the SF-36v2 PCS
• To assess improvement in cardiomyopathy as measured by NT-proBNP	• CFB to Week 50 in NT-proBNP
Exploratory	
• To assess improvement in cardiac amyloidosis as measured by changes in NT-proBNP, cTnT, dFLC, and CRP	• CFBs over time in NT-proBNP, cTnT, dFLC, and CRP
• To assess cardiac response, defined as a decrease of > 30% in NT-proBNP levels and a decrease of NT-proBNP > 300 ng/L	• Proportion of patients with cardiac response at Weeks 14, 26, 38, 50, and 74 • Time to cardiac response
• To assess composite cardiac response when both of the following are satisfied o A decrease of > 30% in NT-proBNP levels and a decrease of NT-proBNP > 300 ng/L o GLS% improvement of ≥ 2%	• Proportion of patients with composite cardiac response at Weeks 14, 26, 50, and 74
• To assess the CCI [redacted]: CCI [redacted]	CCI [redacted]
• To assess graded cardiac response using the following criteria: o Cardiac complete response (CarCR) defined as NT-proBNP ≤ 350 pg/mL (≤ 41.39 pmol/L) o Cardiac very good partial response (CarVGPR) defined as a > 60% reduction in NT-proBNP from the baseline level not meeting CarCR o Cardiac partial response (CarPR) defined as a 31% to 60% reduction in NT-proBNP from the baseline level not meeting CarCR o Cardiac no response (CarNR) defined as a ≤ 30% reduction in NT-proBNP from the baseline level (Muchtar, 2023)	• Proportion of patients with graded cardiac response at Weeks 14, 26, 38, 50, and 74 • Time to graded cardiac responses (CarCR, CarVGPR, and CarPR)
• To assess renal response, defined as a 30% decrease in proteinuria or a drop of	• Proportion of patients with renal response over time

Objectives	Estimand and/or Endpoints
proteinuria below 0.5 g/24 h in the absence of a ≥ 25% decrease in eGFR (Palladini, 2014). Renal response will be assessed in patients with renal involvement at Baseline, defined as those with a baseline urinary protein excretion of more than 0.5 g per day.	• Time to renal response
• To assess renal progression, defined as a ≥ 25% decrease in eGFR, in patients with renal involvement at Baseline	• Time to renal progression
• To assess graded renal response in the absence of renal progression, in patients with renal involvement at Baseline, using the following criteria: o Renal complete response (renCR, 24-hour proteinuria ≤ 200 mg/24 h) o Renal very good partial response (renVGPR, a > 60% reduction in baseline 24-hour proteinuria) o Renal partial response (renPR, a 31% to 60% reduction in baseline 24-hour proteinuria) o Renal no response (renNR, a 30% or less reduction in baseline 24-hour proteinuria) (Muchtar, 2021)	• Proportion of patients with graded renal response (in the absence of renal progression) over time • Time to graded renal response (renCR, renVGPR, and renPR)
• To assess hepatic response in patients with hepatic involvement (defined as ALP ≥ 1.5 ULN) at Baseline, where hepatic response is defined as a decrease of ≥ 50% and/or normalization of serum ALP level (Comenzo, 2012).	• Proportion of patients with hepatic response over time (Muchtar, 2018) at Weeks 14, 26, 38, 50, and 74 • Time to hepatic response
• To assess graded hepatic response in patients with hepatic involvement at Baseline, using the following criteria: o Hepatic complete response (hepCR, nadir ALP ≤ ×2 lower limit of normal) o Hepatic very good partial response (hepVGPR, a > 60% reduction from the baseline ALP level to a nadir above ×2 lower limit of normal) o Hepatic partial response (hepPR, a 31% to 60% reduction from baseline ALP) o Hepatic no response (hepNR, a ≤ 30% reduction from baseline ALP)	• Proportion of patients with graded hepatic response at Weeks 14, 26, 38, 50, and 74 • Time to graded hepatic response (hepCR, hepVGPR, and hepPR)
• To assess changes in effects on liver function including AST, ALT, ALP, and GGT	• CFBs over time in liver function including AST, ALT, ALP, and GGT

Objectives	Estimand and/or Endpoints
• To assess changes in effects on renal function including eGFR, serum creatinine, and 24-hour urine protein	• CFBs over time in renal function including eGFR, serum creatinine, and 24-hour urine protein
• To assess QoL as measured by the KCCQ domain scores, the SF-36v2® scaled domain scores, and EQ-5D-5L™ scores	• CFBs over time in QoL as measured by the KCCQ domain scores and summary scores, SF-36v2 scaled domain scores, and MCS and EQ-5D-5L index score and VAS
• To assess change in NYHA Functional Classification	• CFB in NYHA Functional Classification over time
• To describe the PK profile based on serum levels of CAEL-101	• Serum CAEL-101 concentrations by visit
• To assess the immunogenicity of CAEL-101	• ADAs, incidence, response categories, and titers, as well as NAb incidence and titers

[a] The 5 key secondary endpoints to be tested sequentially are CFB to Week 50 in KCCQ-OS, CFB to Week 50 in 6MWT, Time to ACM, Frequency of CVH and CFB to Week 50 in GLS%

1.2. Study Design

Protocol CAEL101-301 and Protocol CAEL101-302 are double-blind, randomized, multicenter, international Phase 3 studies of CAEL-101 combined with standard-of-care PCD treatment versus placebo combined with standard-of-care PCD treatment in Mayo stage IIIb (CAEL101-301) or IIIa (CAEL101-302) PCD treatment-naïve AL amyloidosis patients. The PETP part of each study will end on the date of the last participant randomized plus 18 months (LPI+18M).

Following screening and randomization, participants will receive blinded treatment of CAEL-101 (1000 mg/m^2) + PCD treatment or placebo + PCD treatment during the PETP.

All participants who completed the PETP will have the option to continue into the Open-Label Extension Period, which consists of a Transition (Loading Dose) Period and an Open-Label Extension Treatment Period. During the Open-label Extension Treatment Period, all participants will receive CAEL-101 in addition to any standard of care prescribed by the physician. Final study assessments will be performed post-treatment in End of Treatment and follow-up visits.

Participants in both study intervention groups will be followed from randomization until death from any cause, heart transplant, LVAD implantation, or the study-level EOS for each study.

For full details on study design, refer to individual study protocols.

2. STATISTICAL HYPOTHESES

The primary objective of pooling these 2 studies is to determine whether CAEL-101 and treatment for PCD improve overall survival and/or decrease the frequency of CVH in Mayo stage IIIa or IIIb AL amyloidosis patients who are treatment naïve, compared to treatment for PCD and placebo.

The primary statistical hypotheses to be tested are as follows:

- H_0: Neither time to ACM nor frequency of CVH is different between CAEL-101 and placebo in participants with Mayo stage IIIa or IIIb AL amyloidosis
- H_a: At least one of time to ACM and frequency of CVH is different between CAEL-101 and placebo in participants with Mayo stage IIIa or IIIb AL amyloidosis

The effects of CAEL-101 on safety and tolerability, while being included as primary objectives, are not to be tested inferentially, and thus, there is no structured hypothesis for those outcomes.

3. SAMPLE SIZE DETERMINATION

The 2 studies were originally designed as event-driven studies and were to continue until the required number of deaths were observed. The power was estimated for the primary endpoint of time to ACM. Sample sizes of 124 and 267 participants, respectively, for CAEL101-301 and CAEL101-302, were randomized in a 2:1 ratio of CAEL-101+PCD:placebo+PCD while 101 and 88 total deaths, respectively, for CAEL101-301 and CAEL101-302, were expected to provide at least CCI power with a statistical significance level of 5% and an assumed hazard ratio of CCI and CCI, respectively, for CAEL101-301 and CAEL101-302.

The primary efficacy endpoint has been revised to a hierarchical combination of time to ACM and frequency of CVH over PETP (CAEL101-301 Protocol Amendment 6.0 and CAEL101-302 Protocol Amendment 7.0) based on the pooled data.

The Finkelstein-Schoenfeld (F-S) test (Finkelstein, 1999) will be used to compare the hierarchical composite endpoint between CAEL101 and placebo. Simulations based on estimates of 12-month survival rates of 0.56 (CAEL101-301) and 0.82 (CAEL101-302) and rates of CVH per participant per year of 0.64 (CAEL101-301) and 0.33 (CAEL101-302) for participants result in greater than CCI power at a 2-sided significance level of 0.05 to test the hypothesis that neither time to ACM nor the cumulative frequency of CVH is different between CAEL-101 and placebo. The simulations also assume an equal treatment effect for both ACM and CVH (eg, hazard ratios of CCI [CAEL101-301] and CCI [CAEL101-302] for time to ACM and rate ratios of CCI (CAEL101-301) and CCI (CAEL101-302) for the CVH).

Due to changes in the SoC and the treatment landscape for AL amyloidosis patients from the time Study CAEL101-301 and Study CAEL101-302 were initiated, there is a possibility that the target treatment effects (based on ACM hazard ratios) may be smaller, but still clinically meaningful. To accommodate this possibility, power calculations were also performed for the treatment effect of hazard ratio = CCI (corresponding to a 30-35% relative risk reduction in ACM at Month 12). By pooling Studies CAEL101-301 and CAEL101-302, the power for a hierarchical combination of time to ACM and frequency of CVH is greater than CCI.

Simulations were also performed to estimate the power for the secondary endpoints, including time to ACM, frequency of CVH, CFB at Week 50 in 6MWT, KCCQ-OS, and GLS%. For time to ACM, the power is greater than CCI based on the log-rank test statistic at a 2-sided significance level of 0.05, assuming a hazard ratio of CCI for Studies CAEL101-301 and CAEL101-302, respectively. For frequency of CVH, the power is greater than CCI based on a negative binomial regression model at a 2-sided significance level of 0.05, assuming a risk ratio of CCI and CCI for the 2 studies, respectively.

For the 3 endpoints (KCCQ-OS, 6MWT, and GLS%) with CFB outcomes, the power is based on the ranked ANCOVA with a 2-sided significance level of 0.05. The ranked ANCOVA accounts for the mortality before 50 weeks based on the assumptions of the treatment effect on mortality mentioned above. The power to detect a difference in CFB at Week 50 in KCCQ-OS is greater than CCI, assuming a mean of -3 units, an SD of 18.5 units for the placebo, and a minimum clinically important difference (MCID) of 5 units of the overall score. The power to detect a difference in CFB at Week 50 in 6MWT is greater than CCI, assuming a mean of -21.15 m, an SD of 20.63 m for the placebo, and an MCID of 33 m. The power to detect a difference in CFB

at Week 50 in GLS% is greater than CCI , assuming a mean of -0.05 units, an SD of 6.60 units for the placebo, and an MCID of 2 units.

4. ANALYSIS SETS

The analysis sets are described in the following table and will include the Screened Set, Intent-to-Treat (ITT) Analysis Set, Per-Protocol (PP) Analysis Set, Pharmacokinetic Analysis Set (PKAS), Safety Analysis Set (SS), and ADA Analysis Set (AAS).

Analysis Set	**Description**
Screened Set	All consented participants.
ITT Analysis Set	All participants who were randomized.
SS	All participants who received at least 1 dose of study intervention. Participants will be analyzed according to the study intervention they actually received, regardless of the study intervention to which they were randomized.
PKAS	All participants who received at least 1 dose of CAEL-101 and had at least 1 measurable serum CAEL-101 concentration.
PP Analysis Set	All participants who met the definition of ITT (see above) and who were without important protocol deviations that would affect the interpretability of the efficacy outcomes (Section 6.2.1).
AAS	All participants who received any study intervention and who, after the first dose, had at least 1 reportable result in the ADA assay. ADA analysis in participants will be conducted based on the actual study intervention that they received.

If not otherwise stated in the respective sections, the statistical analyses will be performed using the analysis sets as follows (Table 1).

Table 1: Analyses Performed for Each Analysis Population

Analyses	**ITT**	**SS**	**PP**	**PKAS**	**AAS**
Disposition	X				
Demographics and baseline characteristics	X	X			
Exposure and compliance		X			
Prior and concomitant medications/therapies and PCD therapies		X			
Efficacy: primary	X		X		
Efficacy: secondary	X		X		
Efficacy: exploratory	X				
Safety analysis		X			
Pharmacokinetic analysis				X	
ADA analysis					X

5. STATISTICAL ANALYSES

5.1. General Considerations

5.1.1. Descriptive Summary Statistics

In general, continuous variables will be summarized using descriptive statistics, that is, displaying the number of participants in the respective analysis population, the number of participants with nonmissing values, mean, SD, minimum, lower quartile, median, upper quartile, IQR, and maximum.

Categorical and ordinal variables will be summarized by using frequency counts and percentages. In addition, the number of participants with missing values will be displayed.

CFB in categorical data will be summarized using shift tables where appropriate.

All descriptive statistics, frequency counts, and percentages will be presented by study intervention group and visit, as appropriate.

5.1.2. Technical Specifications for Derived Variables

5.1.2.1. Definition of Baseline

Baseline is defined as the last nonmissing assessment for a given parameter prior to the first study intervention infusion. If there are multiple assessments prior to the first study intervention infusion, see Section 6.1.2.

5.1.2.2. Derivation of Study Day

The date of the first dose of study intervention will be considered Study Day 1 and the day before the first dose of study intervention will be Study Day -1. Thus, there is no Study Day 0.

Study days will be calculated as follows only when the full assessment date is known (ie, partial dates will have missing study days):

For days on or after the first dose of study intervention:

- Date of Assessment – Date of the First Dose of Study Intervention + 1.

For days before the first dose of study intervention:

- Date of Assessment – Date of the First Dose of Study Intervention.

5.1.3. Analysis Reporting Rules

Means, medians, and quartiles will be presented by 1 additional decimal place, and SD will be presented by 2 additional decimal places than the standard presentation level of the respective participant data. Minimum and maximum values will be presented using the same number of decimal places as the participant data. Percentages will be presented by 1 decimal place if not otherwise stated.

If the number of participants in a category is 0, then the percentage will not be displayed, and only a count of 0 will be shown.

All p-values will be 2-sided, unless otherwise stated, and reported to 3 decimal places (eg, 0.XXX). Values < 0.001 will be reported as < 0.001.

All data collected as part of the study will be reported in data listings, sorted by individual study center, participant, and visit/study day, or as appropriate. In general, all date fields will have an associated study day (eg, for a TEAE start date, a corresponding study day will be presented, with the TEAEs sorted within participant by this start day).

5.1.4. Hypothesis Testing and Significance Level

Unless otherwise specified, all hypothesis testing will be performed using 2-sided tests at the 0.05 level of significance.

Hypothesis testing will be conducted in a prespecified sequential order to control for overall type I error.

A 2-sided hypothesis test (at alpha = 0.05) will be performed for the primary efficacy endpoint. If the p-value is significant, then the following key secondary efficacy endpoints based on pooled data will be tested in the sequence as specified below.

- CFB to Week 50 in the KCCQ-OS
- CFB to Week 50 in the 6MWT
- Time to ACM from randomization to the end of PETP
- Frequency of CVH from randomization to the end of PETP
- CFB to Week 50 in GLS%

If any of the p-values are not significant, then the corresponding null hypothesis will not be rejected. The hypotheses for all the subsequent endpoints will not be rejected. The nominal p-values will be computed for descriptive purpose.

Other secondary endpoints will be tested outside of the sequential approach at the 0.05 level of significance without multiplicity adjustment.

- CFB to Week 50 in the SF-36v2 PCS
- CFB to Week 50 in NT-proBNP

All exploratory efficacy endpoints will be tested at the 0.05 level of significance without multiplicity adjustment.

5.1.5. Analyses for China Registration

To support the registration in China, selected analyses will be performed for China cohort and East Asia cohort, in addition to the subgroup analyses specified in Section 5.7.7.

The China cohort will include all participants enrolled at the sites in China, and the East Asia cohort will include participants from East Asia countries including China, Japan, and South Korea.

Analysis sets for regional cohorts (China and East Asia) will follow the definition in Section 4. The analyses may include, but are not limited to disposition, demographic and baseline

characteristics, extent of exposure, prior and concomitant medications, efficacy, safety, and PK analyses.

All statistical analyses will be considered exploratory in the China cohort or the East Asia cohort. Considering the limited number of participants in the China cohort and the East Asia cohort, only descriptive statistics will be presented for a subset of tables and figures. A demographic summary will be presented by cohort median age for the China cohort and East Asia cohort. No other subgroup analysis will be performed for the China cohort or the East Asia cohort. No separate listings of individual participants in the China cohort or the East Asia cohort will be produced.

5.1.6. Analyses for Japan Registration

The analyses for the Japanese population will be conducted to support the drug application submission in Japan and to assess the consistency in efficacy, PK/PD, and safety between the entire population (or non-Japanese only for PK/PD) and the Japanese population if feasible. These analyses will not be included in the clinical study report but will be included in the Japan CTD.

The following summary outputs will be generated for the subgroup of Japanese population. The analyses will be based on Japanese population, defined as those enrolled from sites in Japan. Unless otherwise specified, the endpoint and variable definitions and methods used for these outputs will be the same as described for overall population.

- Summaries of study disposition, demographics and baseline characteristics, disease characteristics, and prior and concomitant medications
- Summaries of primary and key secondary endpoints in the primary analysis population
- Summaries of PK concentration for the subgroup of Japanese population and the subgroup of non-Japanese population
- Summary of study drug exposure or duration
- Summaries of TEAE overview
- Summary of the following Adverse Events by SOC and PT:
  TEAEs and SAEs, TEAEs by highest severity, TEAEs related to study treatment, TEAEs leading to study drug discontinuation, most common TEAEs, TEAEs by narrow SMQ for hypersensitivity or anaphylactic reaction, TEAEs leading to death, and TEAEs of special interest (if any).

Analyses for other selected endpoints may be performed if deemed necessary. All analyses will be performed descriptively without formal statistical testing.

5.2. Study Participants

The number of screened participants and the number and percentage of participants by each screen failure reason will be reported.

The number and percentage of participants who were randomized, who received randomized treatment, who prematurely discontinued from the study treatment or withdrew early from the study and reasons for discontinuation, and who completed the PETP will be tabulated by study intervention group. Study duration, defined as the time from the informed consent date to the end of PETP or study discontinuation date, whichever occurs first. Study duration will be summarized by study intervention group. Information will be reported using the ITT Analysis Set. It will also be reported by study.

KM (Kaplan, 1958) plots of time to discontinuation from treatment and study will be provided.

Tabulation of the number of participants by region, country, and individual study center will be provided.

5.3. Primary Efficacy Endpoint Analysis

The primary efficacy endpoint will be analyzed based on the ITT Analysis Set on pooled data.

5.3.1. Primary Efficacy Endpoint

The primary efficacy endpoint is the hierarchical combination of time to ACM and frequency of CVH based on pooled data from Protocol CAEL101-301 and Protocol CAEL101-302 up to LPI+18M. Time to ACM is defined as the number of weeks from the date of randomization to the date of death if it is on or before the end of PETP. Otherwise, participants who are alive at the end of PETP (LPI+18M) will be censored at their last known alive date recorded within the PETP. CVH are those events that were adjudicated and confirmed as such by the Clinical Event Adjudication Committee (CEAC).

Participants may be randomized up to a calendar day in advance of the first dose of double-blind study intervention, but, in general, most participants will be treated on the same day as randomization. Participants who prematurely discontinued study treatment or withdrew early from the study will be included in the analysis using the date of death or censoring timepoint (ie, the last known alive date recorded within the PETP), as appropriate. If allowed by country regulatory authorities and/or consented to by the participant, study personnel may use public records to check for death for any participant considered lost to follow-up and for participants who withdrew consent.

5.3.2. Estimand

The primary estimand attributes are as follows:

- Study intervention:
  - CAEL-101 plus standard-of-care PCD treatment
  - Placebo plus standard-of-care PCD treatment
- Population: adults (≥ 18 years of age) with Mayo stage IIIa or IIIb AL amyloidosis
- Variable: a hierarchical combination of time to ACM and frequency of CVH
- Summary measure: win ratio (WR) of CAEL-101 versus placebo for the hierarchical composite of time to ACM and frequency of CVH

- Intercurrent events (ICEs) and corresponding strategies (Section 6.4.1):

ICE	Strategy Addressing the ICE
Heart transplant/LVAD implantation	Composite-variable strategy: participants with ICEs will be treated as having died.
Study intervention discontinuation with reasons other than heart transplant/LVAD implantation	Treatment policy strategy: the ICE will be ignored and the observed data after the ICE are included in the analysis.
Selected important protocol deviations (Section 6.2.1) due to the following: • Intake of prohibited medication • Missing more than 20% of the total planned CAEL-101 doses	Treatment policy strategy: the ICE will be ignored and the observed data after the ICE are included in the analysis.

5.3.3. Main Analytical Approach

Hypothesis testing with the F-S test (Finkelstein, 1999) and treatment effect estimation with the WR method (Pocock, 2012) are based on pairwise comparisons of participant outcomes with the comparisons performed in a hierarchical manner (eg, first compare participant outcomes based on time to ACM and, if there is no winner due to censoring, then compare based on frequency of CVH).

This method compares all pairs of participants, including comparisons with participants belonging to the same treatment group. For any pair of participants, let

$$u_{ij}\ (i \neq j) = \begin{cases} 1, \text{if participant } i \text{ does better than participant } j \text{ (or wins)} \\ 0, \text{if the comparison is undetermined} \\ -1, \text{if participant } i \text{ does worse than participant } j \text{ (or loses)} \end{cases}$$

If $u_{ij} = 1$, then it means that $u_{ji} = -1$. The score for participant i is the sum of the u_{ij} over all pairwise comparisons of participant i with all other participants.

The test is based on the principle that each participant is compared with every other participant within each stratum in a pairwise manner. The method recognizes the highest importance of "death." The pairwise comparison proceeds in a hierarchical fashion using "death" first, assigning + 1 to the "better" participant and - 1 to the "worse" participant.

The F-S scoring algorithm is shown in Figure 1. For a complete list of scenarios, see Table 2.

The stratified F-S test statistic (Finkelstein, 1999) is based on the stratum-specific sum of these scores where the strata are defined by study (CAEL101-301 or CAEL101-302), geographic region (North America or the rest of the world; Section 5.3.3.2) and daratumumab usage (yes or no) at Baseline (-28 ≤ study day ≤ 1), resulting in a total of 8 strata (2 × 2 × 2). If the number of participants per study intervention group in any of these strata is less than 2, the strata of daratumumab usage at baseline may be collapsed, resulting in a total of 4 strata.

The F-S test statistic is based on the sum of the scores for the treated group.

Let $D_i = 1$ for participants in the CAEL-101 group and $D_i = 0$ for participants in the placebo group. Using the u_{ij} for every pair of participants defined above, a score is assigned to each participant.

The participants are divided into 8 strata (k = 1, 2, 3, 4, 5, 6, 7, 8). If A_k is the set of indices of the n_k participants in the k^{th} stratum and U_i is calculated within strata (for $i \in A_k;\ U_i = \sum_{j \in A_k} u_{ij}$), then the test statistic is based on

$$T = \sum_k \sum_{i \in A_k} D_i U_i$$

Let m_k be the total number of participants in the k^{th} stratum who are in the CAEL-101 group. Then, in the k^{th} stratum,

$$E(D_i) = m_k/n_k \text{ and}$$

$$Cov\left(D_i, D_j\right) = \frac{m_k\,(n_k - m_k)}{(n_k - 1)n_k^2}\,(\delta_{ij} n_k - 1)$$

where δ_{ij} is the Kronecker delta (valued as 1 if i=j and 0 otherwise).

As $\sum_{i \in A_k} U_i = 0$, this implies that the mean of T is 0 and its variance is

$$V = \sum_k \frac{m_k(n_k - m_k)}{n_k(n_k - 1)}\,(\sum_{i \in A_k} U_i^2)$$

The hypothesis of interest is tested by comparing $T/\sqrt{V}$ to the normal distribution. The p-value from the F-S test will be presented.

The WR (Pocock, 2012) and its CI will be calculated. For example, WR = 1.2 means that a participant in the treatment group is 20% more likely to "win" than a participant in the placebo group based on time ACM and frequency of CVH, if they are not tied.

The stratified WR method (Dong, 2018) will compare each participant in the CAEL-101 group, with each participant in the placebo group within each stratum, and combine the stratum-specific wins to estimate the stratified WR. The stratified WR will be calculated by adding all the weighted wins from the treatment group and dividing it by all the weighted wins from the placebo group. The weights are chosen as the reciprocal of the stratum size. The SE of log (WR) will be estimated from an asymptotic normal distribution (Dong, 2018). An approximately 95% CI of WR then can be estimated from the CI of log (WR).

Figure 1 Finkelstein-Schoenfeld Scoring Algorithm

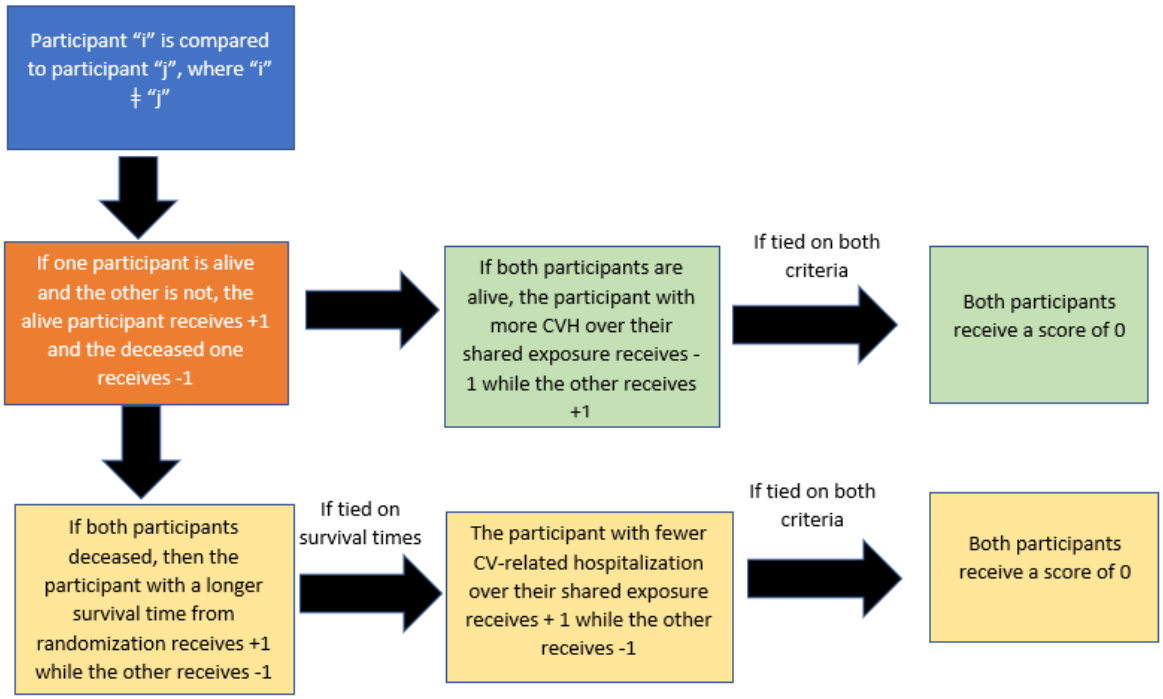


Table 2: Finkelstein-Schoenfeld Scoring Algorithm for the Primary Efficacy Endpoint

Scenario	Participant i/j	Time to ACM		Frequency of CVH[c]	Score
		Vital Status	Survival Time for Deceased /Censoring Time for Alive (from Randomization)		
1	i	Deceased	Low	-	-1
	j	Alive	High	-	+1
2	i	Deceased	Same	-	-1
	j	Alive	Same	-	+1
3[a]	i	Deceased	High	High	-1
	j	Alive	Low	Low	+1
4[a]	i	Deceased	High	Low	+1
	j	Alive	Low	High	-1
5[a]	i	Deceased	High	Tied	0
	j	Alive	Low	Tied	0
6	i	Deceased	Low	-	-1
	j	Deceased	High	-	+1
7[b]	i	Deceased	Tied	High	-1
	j	Deceased	Tied	Low	+1

Scenario	Participant i/j	Time to ACM		Frequency of CVH[c]	Score
		Vital Status	Survival Time for Deceased /Censoring Time for Alive (from Randomization)		
8[b]	i	Deceased	Tied	Tied	0
	j	Deceased	Tied	Tied	0
9[b]	i	Alive	-	High	-1
	j	Alive	-	Low	+1
10[b]	i	Alive	-	Tied	0
	j	Alive	-	Tied	0

[a] Since participant j was alive but censored before participant i was deceased, they will be considered tied with respect to time to ACM and compared by the frequency of CVH.
[b] Participants who are both alive or tied with respect to the time to ACM will be compared by the frequency of CVH.
[c] If the 2 participants have different follow-up times for CVH, the smaller of the 2 follow-up times (shared duration) will be used in comparing the frequency of CVH.

5.3.3.1. Handling of Dropouts or Missing Data

No imputation will be done for unknown vital status or CVH status for the primary efficacy endpoint.

For the component of the primary efficacy endpoint, time to ACM, the date of death will be the key data element used to determine the outcome. As all participants will be followed regardless of whether they continued study intervention or study participation, it is not expected that there will be any missing death dates for those participants who died. Participant data will be censored at the last known alive date recorded. In the event of a participant's death, the complete date of death, including day, month, and year, and the cause of death will be obtained. If the date of death is missing or partially missing, the imputation rules as described in Section 6.1.4 should be applied.

Participants who discontinued study treatment for reasons other than death will be followed for survival status until death from any cause, heart transplant, LVAD implantation, or the study-level EOS. Follow-up visits may be conducted approximately every 4 weeks (± 7 days) following the End of Treatment Visit or as directed by the Sponsor. If allowed by country regulatory authorities and/or consented to by the participant, study personnel may use public records to check for death for any participant considered lost to follow-up and for participants who withdrew consent.

If the admission date of CVH is missing or partially missing, the imputation rules as described in Section 6.1.5 will be applied.

5.3.3.2. Multicenter Studies

The efficacy analyses will be performed using geographic regions (North America or the rest of the world) as the stratification factor. Clinical centers will be categorized into geographic regions based on the country where they are located, as shown in Table 3.

Note that a listing of randomization participant counts by region/country/individual study center will be provided.

Table 3: Geographic Region Definitions

Geographic Region	Country
North America	Canada
	United States
RoW	Australia
	Austria
	Belgium
	Brazil
	China
	Czech Republic
	France
	Germany
	Greece
	Israel
	Italy
	Japan
	Netherlands
	Poland
	Russian Federations
	South Korea
	Spain
	United Kingdom

Note: The final analysis will include all countries other than the United States and Canada in the RoW classification.
Abbreviation: RoW = rest of the world

5.3.4. Sensitivity Analysis

The primary efficacy endpoint analysis will also be performed on the PP Analysis Set.

The ICEs of heart transplant/LVAD implantation are addressed with the composite-variable strategy in the primary efficacy endpoint estimand. As a sensitivity analysis, the primary efficacy endpoint analysis will also be performed where participants with heart transplant/LVAD implantation will be censored at the time of heart transplant/LVAD implantation for ACM.

5.4. Secondary Efficacy Endpoint Analysis

The secondary efficacy endpoints will be analyzed using the ITT Analysis Set.

If the p-value of primary efficacy endpoint testing is significant, then the following key secondary efficacy endpoints will be tested in the sequence as specified below:

- CFB to Week 50 in the KCCQ-OS
- CFB to Week 50 in the 6MWT

- Time to ACM from randomization to the end of PETP
- Frequency of CVH from randomization to the end of PETP
- CFB to Week 50 in GLS%

Other secondary endpoints that will be tested outside of the sequential approach are as follows:

- CFB to Week 50 in the SF-36v2 PCS
- CFB to Week 50 in NT-proBNP

Details of each secondary efficacy endpoint are defined as follows.

CFB to Week 50 in the KCCQ-OS

The Kansas City Cardiomyopathy Questionnaire (KCCQ) is a self-administered questionnaire utilized to collect participant's QoL assessment. It is a valid, reliable, and responsive health-related QoL instrument for participants with congestive heart failure and may serve as a clinically meaningful outcome measure in cardiovascular research, participant management, and quality assessment. The KCCQ has also been assessed via the Medical Device Development Tool qualification process, where it was concluded there was sufficient evidence to support the validity and reliability of the KCCQ for the qualified context of use (FDA, 2016).

The KCCQ includes 23 items that map to 7 domains: physical limitations, symptom stability, symptom frequency, symptom burden, self-efficacy, QoL, and social limitations. To facilitate interpretation, all domain scores are represented on a 0-to-100 scale, where lower scores represent more severe symptoms and/or limitations and a score of 100 indicates no symptoms, no limitations, and excellent QoL.

The KCCQ-OS provides an overall score for participant's QoL (Green, 2000). It is calculated as the mean of the physical limitation score, total symptom score, QoL score, and social limitation score. See Section 6.3.1 for scoring calculation details for all domains and the KCCQ-OS.

Time to ACM from Randomization to the End of PETP

Time to ACM will be defined as the number of weeks from the date of randomization to the date of death if it is on or before the end of PETP. Participants who are alive at the end of PETP will be censored at their last known alive date recorded within the PETP.

Time to event or censoring in weeks will be defined from randomization to the date of event or censoring + 1 day, divided by 7. The date of event is the date of death. The date of censoring is a participant's last known alive date recorded.

Participants who prematurely discontinued study treatment or withdrew early from the study will be included in the analysis (even if after discontinuation of study treatment or early withdrawal from the study) using the date of death or censoring timepoint (ie, the last known alive date recorded within the PETP), as appropriate.

CFB to Week 50 in the 6MWT

The 6MWT measures the distance a participant can walk on a flat, hard surface in a period of 6 minutes. It evaluates the global and integrated response of all the systems involved during exercise. This is a self-paced test; participants choose their own intensity of exercise and are

allowed to stop and rest during the test (Maurer, 2018). The 6MWT is performed according to 6MWT Manual, which is reviewed by the sites.

The 6MWT has been utilized as a measure of amyloidosis progression for other indications.

CFB to Week 50 in the GLS%

Participants will undergo echocardiography for the assessment of GLS% according to the Schedule of Assessments (SoA) of each study's protocol. GLS% expresses longitudinal shortening as a percentage (change in length as a proportion to baseline length) and is reported as a negative value. A more negative value is usually associated with cardiac improvement. The American College of Cardiology suggests that GLS% $\geq$ -16% is abnormal, GLS% $\leq$ -18% is normal, and GLS% values between -16% and -18% are borderline (Yang, 2018).

Frequency of CVH from Randomization to the End of PETP

Hospitalization information will be collected for participants during the study. Whether or not such hospitalizations are CVH related will be reviewed and adjudicated by CEAC according to definitions and processes specified in the CEAC charter (Protocol Section 6.15). Only adjudicated CVH will be included in the analysis of this secondary endpoint. Frequency of CVH will be assessed from randomization to the end of PETP.

CFB to Week 50 in the SF-36v2 PCS

The SF-36v2 is a self-administered questionnaire containing 36 items that measure health on functional status, well-being, and overall evaluation of health in 8 domains.

Except for the 1 single item of self-evaluated transition (Item 2), the scores of the other 35 items are summated into 8 multi-item scales, including vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social role functioning, and mental health. The 8 multi-item scales are aggregated into PCS and MCS.

Norm-based T scores will be obtained for each of the above domain scales, PCS, and MCS using QualityMetric Health Outcomes Scoring Software Version 2.0.7653.27808.

For all scales and summary components, higher scores demonstrate better QoL. CFB to Week 50 in the PCS norm-based T scores is one of the key secondary endpoints in this study. CFB to Week 50 for the domain scales and MCS is defined as exploratory endpoints (Section 5.5.1).

CFB to Week 50 in NT-proBNP

To assess improvement in cardiomyopathy as measured by NT-proBNP, CFB to Week 50 in NT-proBNP will be evaluated.

5.4.1. Estimand

5.4.1.1. Time to ACM

The estimand attributes for time to ACM endpoint are as follows:

- Study intervention:
  - CAEL-101 plus standard-of-care PCD treatment
  - Placebo plus standard-of-care PCD treatment

- Population: adults (≥ 18 years of age) with Mayo stage IIIa or IIIb AL amyloidosis
- Variable: time to ACM from randomization up to the date of death (for participants who died) or the end of PETP
- Summary measure: hazard ratio of CAEL-101 versus placebo
- ICEs and corresponding strategies (Section 6.4.1):

ICE	Strategy Addressing the ICE
Heart transplant/LVAD implantation	Composite-variable strategy: participants with ICEs will be treated as having died.
Study intervention discontinuation with reasons other than heart transplant/LVAD implantation	Treatment policy strategy: the ICE will be ignored, and the observed data after the ICE are included in the analysis.
Selected important protocol deviations (Section 6.2.1) due to the following: • Intake of prohibited medication • Missing more than 20% of the total planned CAEL-101 doses	Treatment policy strategy: the ICE will be ignored, and the observed data after the ICE are included in the analysis.

5.4.1.2. Frequency of CVH

The estimand attributes for frequency of CVH endpoint are as follows:

- Study intervention:
  - CAEL-101 plus standard-of-care PCD treatment
  - Placebo plus standard-of-care PCD treatment
- Population: adults (≥ 18 years of age) with Mayo stage IIIa or IIIb AL amyloidosis
- Variable: frequency of CVH from randomization to the end of PETP
- Summary measure: incident risk ratio of CAEL-101 versus placebo
- ICEs and corresponding strategies (Section 6.4.1):

ICE	Strategy Addressing the ICE
Study intervention discontinuation	While-on-treatment strategy: the observed values up to the ICE are included in the analysis.
Selected important protocol deviations (Section 6.2.1) due to the following: • Intake of prohibited medication • Missing more than 20% of the total planned CAEL-101 doses	Treatment policy strategy: the ICE will be ignored, and the observed data after the ICE are included in the analysis.

5.4.1.3. CFB to Week 50

The estimand attributes for CFB to Week 50 are as follows:

- Study intervention:

  - CAEL-101 plus standard-of-care PCD treatment
  - Placebo plus standard-of-care PCD treatment
- Population: adults (≥ 18 years of age) with Mayo stage IIIa or IIIb AL amyloidosis
- Variable: CFB to Week 50 in the NT-proBNP, KCCQ-OS, 6MWT, SF-36v2 PCS, and GLS%
- Summary measure: Hodges-Lehmann estimate of the median difference between CAEL-101 and placebo for CFB to Week 50 in NT-proBNP, KCCQ-OS, GLS%, 6MWT and SF36v2 PCS
- ICEs and corresponding strategies (Section 6.4.1):

ICE	Strategy Addressing the ICE
Death or heart transplant or LVAD implantation	Composite-variable strategy: the ICE is incorporated in the rank calculation.
Study intervention discontinuation other than death, heart transplant/LVAD	Composite-variable strategy: the ICE is incorporated in the rank calculation.
Selected important protocol deviations (Section 6.2.1) due to the following: • Intake of prohibited medication • Missing more than 20% of the total planned CAEL-101 doses	Treatment policy: the ICE will be ignored, and the observed data after the ICE are used in the rank calculation and included in the analysis.

5.4.2. Main Analytical Approach

5.4.2.1. Time-to-Event Analysis

Time-to-event methods will be used to assess ACM. The time to ACM will be analyzed as follows:

- Testing: A log-rank statistic stratified by study (CAEL101-301 or CAEL101-302), geographic region (North America or the rest of the world; Section 5.3.3.2) and daratumumab use (yes or no) at Baseline and will be provided.
- Estimation: The hazard ratio of CAEL-101 versus placebo based on a Cox proportional hazard model (Cox, 1972) stratified by study and geographic region, and adjusted by daratumumab usage at Baseline will be provided.

Additionally, the KM estimates (95% CI) at Week 50 will be presented. The 95% CI for KM estimates will be calculated using the log-log transformation method.

The corresponding Relative Risk Reduction (RRR) at Week 50 will also be presented.

A KM plot will also be provided by the study intervention groups.

The proportional hazard assumption made in the Cox proportional hazard model analysis may be assessed by graphical methods such as log (-log) plots.

If there is evidence of deviation from the proportionality hazards assumptions that impact the interpretation of the conclusions from the Cox model, alternative approaches may be explored (eg, restricted mean survival time).

5.4.2.2. Negative Binomial Model

Frequency of CEAC-adjudicated CVH will be analyzed using negative binomial regression analysis with study intervention, study, geographic region, daratumumab usage at Baseline, and offset term equal to log of each participant's follow-up time for CVH included in the model.

The estimated incidence risk ratio between study intervention, the 95% CI, and the p-value, as well as the frequency of CVHs per year and corresponding 95% CI will be reported.

5.4.2.3. Ranked ANCOVA

CFB to Week 50 in NT-proBNP, KCCQ-OS, GLS%, 6MWT, and SF-36v2 PCS will be analyzed using a ranked ANCOVA model (Lachin, 1999). The ranked ANCOVA model will include treatment factor, study, geographic region and daratumumab usage at Baseline as fixed effects. The ranked baseline value corresponding to each endpoint will be included as a covariate. The p-value of the estimated LS-mean difference between the 2 intervention groups from the ranked ANCOVA model will be provided.

For the ranked ANCOVA model, participants will be ranked from the worst to the best as follows:

- Participants who have missing values at Week 50 and had any of these events: died / had a heart transplant / LVAD implantation, prior to Week 50, earlier event will be ranked worse than later such events.
- Next, participants who have missing values at Week 50 and whose vital status is unknown at Week 50 will be ranked according to time on study intervention from the worst (shortest time) to the best (longest time).
- Next, participants who have missing values at Week 50 and are known to be alive at Week 50 will be ranked according to time on study intervention from the worst (shortest time) to the best (longest time) up to Week 50.
- Next, participants with available data at Week 50 will be ranked from the worst to the best based on their values (eg, lower scores in KCCQ-OS represent poorer QoL and should be ranked worse than the higher scores in KCCQ-OS).

If a participant falls in multiple categories as described above, the worst scenario will be used to determine the ranking. Details of obtaining the ranks for participants are described in Section 6.4.2.

In addition, the Hodges-Lehmann estimate of the median treatment difference with associated 95% CI will also be presented based on the observed data.

Descriptive statistics including the number of participants, mean, SD, SE, 95% CI, median, range, and IQR by visit and intervention group will be provided for the actual values and CFB for each secondary endpoint. The graphs of means over time (absolute and CFB values) by study intervention group will be provided.

5.4.3. Handling of Dropouts or Missing Data for Secondary Efficacy Endpoints

As described in the protocol, for participants who discontinued study intervention, all attempts will be made to collect their data until EOS.

In case there are still missing data for the secondary endpoints even after every attempt has been made to recover the missing information, they will be handled as follows:

- Time to ACM: refer to Section 5.3.3.1 for details.
- Frequency of CVH: no imputation will be done for the negative binomial analysis (Section 5.4.2.2).
- CFB to Week 50 endpoints (NT-proBNP, KCCQ-OS, 6MWT, SF-36v2 PCS, and GLS%): no imputation will be done for the primary analysis using ranked ANCOVA (Section 5.4.2.3). For missing data handling for the sensitivity analysis, see Section 5.4.4.

5.4.4. Sensitivity Analysis

The analyses of the secondary efficacy endpoints (Section 5.4.2) will also be performed on the PP Analysis Set.

In addition, the following analyses will be performed on the ITT Analysis Set.

- For time to ACM, the sensitivity analysis will also be performed where participants with heart transplant/LVAD implantation will be censored at the time of heart transplant/LVAD implantation. In addition, other covariates (sex, age, race, and free light-chain [FLC] subtype) will be included in the Cox model (Section 5.4.2.1).
- The CFB of secondary efficacy endpoints (NT-proBNP, KCCQ-OS, GLS%, 6MWT, and SF-36v2 PCS) at Week 50 will be evaluated using an ANCOVA model with missing data imputed. The ANCOVA model will include treatment factor, baseline value (ie, baseline NT-proBNP, baseline 6MWT, baseline KCCQ-OS, baseline SF-36v2 PCS, and baseline GLS% as corresponding to each endpoint), study, geographic region and daratumumab usage at baseline as covariates. Missing data for participants in both study intervention groups at Week 50 will be imputed using random sampling (with replacement) from the bottom 10% of the observed CFB values of the placebo participants at Week 50, with the imputed values capped at each participant's worst possible CFB values. The imputation will be done separately for each study. A total of 100 imputed datasets will be created and analyzed using the ANCOVA approach. The 100 resulting treatment effect parameters and SE will be combined using Rubin's rule (Rubin, 1987). The estimated LS means (± SE) by each study intervention group as well as the difference in LS means between the 2 study intervention groups along with the p-value will be provided.

5.5. Exploratory Analysis

Exploratory endpoints may be reported separately from the primary clinical study report.

5.5.1. Exploratory Endpoints

5.5.1.1. Patient-Reported Outcome

CFB over time in the following patient-reported outcomes will be evaluated.

- KCCQ domain scores
  - Physical limitations
  - Symptom stability
  - Symptom frequency
  - Symptom burden
  - Self-efficacy
  - QoL
  - Social limitations
- KCCQ summary scores
  - Total symptom score (symptom frequency score and symptom burden score)
  - Clinical summary score (physical limitation score and total symptom score)
- SF-36v2 scaled domain scores
  - Vitality
  - Physical functioning
  - Bodily pain
  - General health perceptions
  - Physical role functioning
  - Emotional role functioning
  - Social role functioning
  - Mental health
- SF-36v2 MCS
- EQ-5D-5L (refer to Section 6.3.2 for derivation):
  - Index score
  - Visual analog score

5.5.1.2. Cardiac Endpoints

- CFB over time in cardiac and amyloidosis biomarkers, including the following:

  - NT-proBNP
  - Highly sensitive cTnT
  - dFLC

    dFLC is the difference between involved and uninvolved FLCs, that is, dFLC = involved FLC – uninvolved FLC. Involved FLC is classified as Kappa or Lambda (whichever is higher at Baseline). Uninvolved FLC is classified as Kappa or Lambda (whichever is lower at Baseline). These classifications will not change. For example, if a participant has a higher Kappa value compared to Lambda at Baseline, the dFLC for this participant will be calculated by (Kappa value – Lambda value) for all other timepoints.
  - CRP
- Proportion of cardiac response at Weeks 14, 26, 38, 50, and 74, where cardiac response is defined as a decrease of > 30% in NT-proBNP levels AND a decrease of NT-proBNP > 300 ng/L
- Proportion of composite cardiac response at Weeks 14, 26, 38, 50, and 74, where composite cardiac response is defined when both of the following criteria are satisfied:
  - A decrease of > 30% in NT-proBNP levels AND a decrease of NT-proBNP > 300 ng/L
  - GLS% improvement of ≥ 2%
- CCI [redacted]
- Proportion of graded cardiac response at Weeks 14, 26, 38, 50, and 74, where graded cardiac response is determined using the following definitions:
  - CarCR defined as NT-proBNP ≤ 350 pg/mL (≤ 41.39 pmol/L)
  - CarVGPR defined as a > 60% reduction in NT-proBNP from the baseline level not meeting CarCR
  - CarPR defined as a 31% to 60% reduction in NT-proBNP from the baseline level not meeting CarCR
  - CarNR defined as a ≤ 30% reduction in NT-proBNP from the baseline level
- Time to cardiac response
- Time to CarCR
- Time to CarVGPR or better

- Time to CarPR or better

Cardiac response, composite cardiac response, and graded cardiac response will be assessed in participants with cardiac involvement, defined as those with a baseline NT-proBNP value ≥ 650 ng/L.

- CFB in NYHA Functional Classification (see Protocol CAEL101-301 Section 8.10 and CAEL101-302 Protocol Section 8.11).

5.5.1.3. Renal Endpoints

- CFB over time in renal function:
  - eGFR (refer to Section 6.1.6 for derivation)
  - Serum creatinine
  - 24-hour urine protein
- Proportion of renal response, which is defined as a 30% decrease in proteinuria or a drop of proteinuria below 0.5 g/24 h in the absence of a ≥ 25% decrease in eGFR (Palladini, 2014). Renal response will be assessed in participants with renal involvement at Baseline, defined as those with a baseline urinary protein excretion of more than 0.5 g per day.
- Proportion of graded renal response, which is determined using the following criteria in the absence of renal progression (a ≥ 25% decrease in eGFR). Similar to renal response, graded renal response will be assessed in participants with renal involvement at Baseline.
  - renCR defined as 24-hour proteinuria ≤ 200 mg/24 h
  - renVGPR defined as a > 60% reduction in baseline 24-hour proteinuria and not meeting renCR
  - renPR defined as a 31% to 60% reduction in baseline 24-hour proteinuria and not meeting renCR
  - renNR defined as a 30% or less reduction in baseline 24-h proteinuria
- Time to renal response
- Time to renCR
- Time to renVGPR or better
- Time to renPR or better
- Time to renal progression, defined as a ≥ 25% decrease in eGFR, in participants with renal involvement at Baseline
- CFB in eGFR categories in chronic kidney disease (Section 6.1.6)

5.5.1.4. Hepatic Endpoints

- CFB over time in liver function:

  - AST
  - ALT
  - ALP
  - GGT
- Proportion of hepatic response at Weeks 14, 26, 38, 50, and 74, where hepatic response is defined as a decrease of ≥ 50% or normalization of the serum ALP level. Hepatic response will be assessed in participants with hepatic involvement (defined as ALP ≥ 1.5 ULN).
- Proportion of graded hepatic response at Weeks 14, 26, 38, 50, and 74, where graded hepatic response is defined using the following criteria in participants with hepatic involvement.
  - Hepatic complete response (hepCR, ALP ≤ ×2 lower limit of normal)
  - Hepatic very good partial response (hepVGPR, a > 60% reduction from the baseline ALP level to a nadir above ×2 lower limit of normal)
  - Hepatic partial response (hepPR, a 31% to 60% reduction from baseline ALP)
  - Hepatic no response (hepNR, a ≤ 30% reduction from baseline ALP)
- Time to hepatic response
- Time to hepatic hepCR
- Time to hepatic hepVGPR or better
- Time to hepatic hepPR or better

5.5.2. Main Analytical Approach

5.5.2.1. Descriptive Summary Statistics

Descriptive statistics, including the number of participants, mean, SD, SE, 95% CI, median, range, and IQR by visit and intervention group, will be provided for the actual values and CFB for the continuous exploratory endpoints.

5.5.2.2. Proportion Comparison

The proportion of participants with cardiac responses, graded cardiac responses, and composite cardiac responses will be provided by study intervention group. The difference in proportion of participants with cardiac response, graded cardiac responses, or composite cardiac responses between the 2 intervention groups, p-value as well as the 2-sided 95% CI will be estimated using the Cochran-Mantel-Haenszel method adjusted by 3 stratification variables: study, geographic region and daratumumab usage at Baseline. No multiple imputation will be done for the missing data. If certain strata are empty or extremely sparse, clinically justifiable collapsing of those strata may be considered. Any such modifications to achieve adequate counts will be documented.

Renal response, graded renal responses, renal progression, hepatic response, and graded hepatic response will be analyzed in a similar fashion. Neither renal response nor renal progression will be categorized as renal stability and will be included in the analysis.

5.5.2.3. Finkelstein-Schoenfeld Method and Win Ratio

CCI

[redacted]

5.5.2.4. Shift Table

CFB in NYHA Functional Classification over time will be summarized by a shift table.

CFB in eGFR categorization (Section 6.1.6) over time will be summarized by a shift table.

5.5.2.5. Time to Response Summary

KM plots by study intervention group will be provided for time to cardiac response, time to graded cardiac responses, time to renal response, time to graded renal responses, time to renal progression, time to hepatic response, and time to graded hepatic responses. Median time to response from KM method may also be reported.

Participants who never reach a response as described above will be censored at the last assessment date.

5.5.3. PK Analyses

For the PKAS, CAEL-101 serum concentrations will be listed and summarized by scheduled sampling time and will include descriptive statistics (N, mean, SD, percentage of the coefficient of variation [%CV], median, minimum, maximum, geometric mean, and geometric %CV). Individual and mean serum concentration-time profile plots will be displayed using scheduled sampling time.

A separate population PK analysis plan will be written for the integrated population PK or PK/PD analyses that could include PK and/or PD data across the clinical development of CAEL-101.

5.5.4. Immunogenicity

Samples for immunogenicity testing will be collected according to the SoA and prior to infusion of study intervention. Samples collected at designated time points will be analyzed using validated ADA assays in all participants included in the AAS. Samples positive in the ADA assay will be further characterized in the ADA titer and the NAb assays.

The summaries of antidrug antibody (ADA) incidence over the duration of the study will include the following response categories:

- ADA negative: An ADA-negative signal in the ADA assay at all timepoints collected for ADA analysis
- ADA positive: An ADA-positive signal in the ADA assay at any timepoint collected for ADA analysis
- Participants who are ADA positive may be further categorized into ADA response categories as follows:
    - Pre-existing immunoreactivity: An ADA-positive response with either of the following 2 conditions met:
    - ADA-positive response at Baseline with all post-first dose ADA results negative, OR
    - ADA-positive response at Baseline with all post-first dose ADA responses less than 4-fold over the baseline titer level
    - Treatment-emergent ADA responses: An ADA-positive response post-first dose when baseline results are negative or missing
    - Treatment-boosted ADA responses: An ADA-positive response post-first dose that is ≥ 4-fold over the baseline titer level when the baseline result is positive

Treatment-emergent or treatment-boosted ADA responses may be further categorized as follows:

- Persistent: ADA responses with 2 or more consecutive ADA-positive samples separated by at least a 16-week period, with no ADA-negative samples in between, irrespective of missing samples
- Indeterminate: ADA-positive sample only at the last collected sample
- Transient: ADA response that is neither a persistent nor an indeterminate response

ADA-positive samples will be further characterized for neutralizing activity in the NAb assay. NAb status categories for a participant are defined as follows:

- NAb positive: A NAb-positive signal in the NAb assay at any timepoint collected for NAb analysis
- NAb negative: A NAb-negative signal in the NAb assay at any timepoint collected for NAb analysis

The following analyses will be generated based on pooled data as well as by study (CAEL101-301 and CAEL101-302).

The incidence of ADA response categories will be summarized as absolute occurrence (n) and percentage (%) of all participants by study intervention. NAb-positive and NAb-negative participants will be summarized under treatment-emergent ADA response and treatment-boosted ADA response categories as absolute occurrence (n) and percentage (%) of all participants by study intervention.

ADA positive will be summarized as absolute occurrence (n) and percentage (%) of all participants by timepoint. ADA titer will also be descriptively summarized by timepoint.

Associations between ADA response categories and systemic exposure to CAEL-101 may be explored for CAEL-101-treated (not placebo) participants. Spaghetti plots of PK concentration-time profiles by ADA response category may be provided for analyzing the potential impact of immunogenicity on PK profiles. Plots of PK parameter values (C_{trough}) and titer values (double y-axis) across time by ADA response category may be provided.

The impact of ADA on the PK analysis will be described as part of the population PK analysis plan and, thus, is not included in this current SAP.

For analyses of the association of immunogenicity with impacts on safety:

- Associations between ADA response categories and serious and severe AEs may be explored, including serious adverse events (SAEs) such as systemic hypersensitivity, anaphylaxis, infusion site reactions lasting more than 24 hours, and other immune-related SAEs.

For analyses of the association of immunogenicity with impacts on efficacy:

- Associations between ADA response categories and key efficacy endpoints or variables may be explored for CAEL-101-treated (not placebo) participants. Plots of key efficacy endpoints may be analyzed for potential impact of immunogenicity on drug efficacy.

5.6. Safety Analyses

The analysis will be descriptive in nature, and no formal statistical tests will be performed. All safety data will be listed and tabulated by study intervention group. Safety data include the following:

- TEAEs
- TESAEs

- Clinical laboratory tests (hematology, chemistry, urinalysis, pregnancy, and special laboratory tests)
- Vital signs, including respiratory rate, heart rate, BP, weight, and temperature
- Twelve-lead ECG, including rate, rate ranges, and diagnostic statements

AEs will be coded using the MedDRA coding dictionary Version 23.0 or higher and presented by MedDRA SOC and PT. Participant incidence of each SOC and unique PT will be tabulated, including for all TEAEs, TESAEs, at least possibly related TEAEs (thus, TEAEs will be tabulated as unrelated versus possibly related/related combined, which will be categorized as "related"), at least possibly related SAEs, TEAEs by severity (National Cancer Institute (NCI)-Common Terminology Criteria for Adverse Events [CTCAE] Grades 1 to 5), TEAEs resulting in discontinuation of study treatment, and TEAEs resulting in death. Note that only AEs that occur on or after the date of signing the informed consent form are entered into the clinical database. Treatment emergent is defined as follows: any AE that starts between the start of the first infusion of study intervention and up to the last study intervention date + 140 days.

Analyses of laboratory, vital sign, and ECG outcomes will include mean CFB, as well as shifts (eg, normal to abnormal) from Baseline.

Graphics of changes over time in safety parameters will be generated, where appropriate.

5.6.1. Extent of Exposure

The SS will be used to summarize study treatment duration, study intervention exposure, and treatment compliance.

Study treatment duration of exposure will be calculated as the date of the last study intervention minus the date of the first study intervention + 1 day. Exposure will be summarized by descriptive statistics and by categories of < 3 months, 3 to < 6 months, 6 to < 9 months, 9 to < 12 months, 12 to < 15 months, 15 to < 18 months, 18 to < 24 months, 24 to < 36 months, and ≥ 36 months.

The number (%) of participants by exposure duration will be tabulated. Tabulation of the number of participants by region and country, by age group (below study median age and above or equal to study median age), and by sex will be provided.

Participants randomized to receive CAEL-101 will receive 1000 mg/m^2 by infusion. The planned total dose will be based on the participant's body surface area (BSA) in square meters, which is calculated using the height and weight obtained during the Screening Period.

The actual dose (milligrams) for an infusion will be calculated using planned volume administered (milliliters), actual volume administered, and planned dose (milligrams) as follow:

$$Actual\ dose\ (mg) = \frac{Actual\ volume\ administered\ (mL)}{Planned\ volume\ administered\ (mL)} \times Planned\ dose\ (mg)$$

The overall compliance of CAEL-101 is defined as follows:

$$Overall\ compliance\ (\%) = \frac{Actual\ total\ dose\ administered\ (mg)}{Planned\ total\ dose\ (mg)} \times 100$$

Treatment compliance will be summarized by descriptive statistics and as frequency counts and percentages for the following categories (< 80% versus ≥ 80%).

Note that PCD is a concomitant treatment given in conjunction with CAEL-101 (or placebo). An exploration of the types of PCD treatments given (cyclophosphamide-bortezomib-dexamethasone, daratumumab, etc.) will be provided via the number (%) of participants receiving each during treatment, and the duration of treatment with each PCD (where sufficient number of participants allows for determination of this) will be summarized similarly, as appropriate.

5.6.2. Adverse Events

An AE is defined as any untoward medical occurrence in a participant administered a pharmaceutical product, at any dose, that is not necessarily related to the treatment (see the study protocol for a more comprehensive description of AEs). The focus of analysis in this study will be on the incidence of TEAEs.

Percentages of participants are based on the total number of treated participants in the particular treatment group.

Related AEs are defined as possibly related or related.

TEAE and TESAE incidence will be summarized by sex, race, study, age group, and NYHA Functional Classification as defined in Section 5.7.7.

5.6.2.1. Overall Summary of AEs

A summary of the incidence of TEAEs will be summarized by unique participant count and percentage per study intervention group including the following:

- Any TEAE
- Any serious TEAE
- Participants with any TEAE leading to discontinuation of study intervention
- Participants with any TEAE resulting in an outcome of death as recorded in the AE case report form page
- TEAE by severity (Grade 1 to Grade 5)
- At least one Grade 3 (or higher) TEAE
- TEAE by relationship (related or not related) to study intervention
- TESAE by relationship (related or not related) to study intervention

Additional summary tables by sex, race, age group, study, and NYHA Functional Classification (Section 5.7.7) will be presented.

5.6.2.2. TEAEs and SAEs by SOC and PT

The number and percentage of participants with TEAEs by study intervention groups and overall will be presented by SOC and PT. Participants are counted once in each SOC and PT.

Percentages will be based on the total number of treated participants in each group in the SS. SOCs and PT will be listed in the order of decreasing frequency of occurrence.

TESAEs will be summarized similarly.

More frequent TEAEs (occurring in at least 25% of participants in the CAEL-101 group) will also be tabulated.

5.6.2.3. TEAEs and TESAEs by SOC, PT, and Relationship

The number and percentage of participants with TEAE will be presented by SOC and PT by relationship. If a participant has more than 1 occurrence of a TEAE, the highest relationship to study treatment will be summarized by SOC and PT.

TESAEs will be summarized similarly.

5.6.2.4. TEAEs and TESAEs by SOC, PT, and Severity

The number and percentage of participants with TEAEs will be presented by SOC and PT and by severity (Grade 1 to Grade 5). If a participant has more than 1 occurrence of a TEAE, the most severe occurrence will be summarized.

TESAEs will be summarized similarly.

5.6.2.5. Deaths, Other TESAEs, and Other Significant AEs

The number of TEAEs leading to discontinuation from the study intervention and leading to death and the number and percentage of participants with the events will be summarized by SOC and PT. Additional summary tables by study will be generated.

5.6.2.6. Infusion-Related Reactions

Infusion reactions will be assessed using the Standardized MedDRA Query (SMQ) for hypersensitivity and anaphylactic reactions. Overall summaries, summaries by SOC and PT, TESAEs, and summaries by relationship and by severity will be provided for the below categories.

- Narrow SMQ for “Hypersensitivity” or “Anaphylactic reaction”
  - Narrow SMQ for “Hypersensitivity” or “Anaphylactic reaction” that occurred within 24 hours of drug administration
- Narrow SMQ for “Hypersensitivity” or “Anaphylactic reaction” excluding the High Level Term (HLT) of Infusion site reactions and administration site reactions NEC
  - Narrow SMQ for “Hypersensitivity” or “Anaphylactic reaction” excluding HLT of Infusion site reactions and administration site reactions NEC that occurred within 24 hours of drug administration

For time period calculation (within 24 hours), if the AE start time is missing or the infusion time is missing, then it will be defined as infusion start date ≤ AE start date ≤ (infusion start date + 1) per participant and infusion start date.

5.6.3. Additional Safety Assessments

5.6.3.1. Vital Signs

Vital signs include HR, respiratory rate, systolic and diastolic BP, and temperature.

CFB in vital signs (systolic BP, diastolic BP, HR, respiratory rate, and temperature) at each visit will be summarized descriptively by study intervention as observed and CFB values. A listing of vital signs will be presented by study intervention group, participant, vital sign, and visit.

5.6.3.2. Laboratory Assessments

Samples for clinical laboratory tests will be obtained according to the SoA and sent to the central laboratory. Results from a local laboratory can be used if central laboratory results are not available for a visit. Clinical laboratory tests that will be performed are listed in Table 4.

Table 4: Clinical Laboratory Tests

Hematology	Chemistry	Urinalysis
• Hematocrit • Hemoglobin • Red blood cells • White blood cells • Platelet count • Neutrophils (absolute) • Lymphocytes (absolute) • Monocytes (absolute) • Eosinophils (absolute) • Basophils (absolute) • MCH • MCHC • MCV • RDW • Reticulocyte count	• AST • ALT • ALP (ALP isozyme, if applicable) • GGT • Albumin • Globulin • Creatinine • Glucose • Total protein • Total bilirubin • Sodium • Potassium • Chloride • CO_2 • Calcium • Phosphorus • Magnesium • Blood urea nitrogen • eGFR • Uric acid	• Specific gravity • pH • Glucose • Protein • Bilirubin • Ketones • Leukocytes • Blood
		Pregnancy
		• Serum hCG • Urine pregnancy test
		Special
		• Serum FLC • Serum and urine immunofixation • iFLC or dFLC or SPEP m-spike • Troponin • NT-proBNP • CRP

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; CO_2 = carbon dioxide; CRP = C-reactive protein; dFLC = involved/uninvolved free light-chain difference; eGFR = estimated glomerular filtration rate; FLC = free light chain; GGT = gamma-glutamyl transferase; hCG = human chorionic gonadotropin; iFLC = involved free light chain; MCH = mean corpuscular hemoglobin; MCHC = mean corpuscular hemoglobin concentration; MCV = mean corpuscular volume; NT-proBNP = N-terminal pro-B-type natriuretic peptide; RDW = red cell distribution width; SPEP = serum protein electrophoresis

Descriptive statistics by time of assessment and study intervention will be presented for each numeric laboratory parameter. CFB, as well as shift tables, will be presented. Laboratory values will be classified as normal, below normal (low), or above normal (high) based on normal ranges supplied by the central laboratory or a local laboratory, if applicable. Shift table will display the

number of participants who have low, normal, or high laboratory values at Baseline cross-tabulated by the worst postbaseline shifts. If a participant's baseline laboratory values are missing, this participant will not be included in the CFB analysis. If a laboratory parameter has an ordinal category defined (eg, urinalysis parameters), number and percentage of participants will be summarized by category and by visit and study intervention. For purposes of analyses, laboratory test results based on standardized units will be used.

Laboratory chemistry values will be categorized into the following 3 classifications. The analysis described above will be done separately for each category by each laboratory parameter.

Laboratory Classification	**Laboratory Parameters**
Liver function/hepatic parameters	AST, ALT, ALP, GGT, total bilirubin, albumin, and total protein
Kidney function/renal parameters	Creatinine and eGFR
Other	Globulin, glucose, sodium, potassium, chloride, CO_2, calcium, phosphorus, magnesium, blood urea nitrogen, and uric acid

Longitudinal plots (absolute values or CFB) may be generated for selected clinical chemistry, hematology, urinalysis, special biomarkers, vital signs, and ECG values.

Laboratory test results will be graded using NCI CTCAE V5.0 toxicity grading. The incidence of the highest toxicity grade (using NCI CTCAE V5.0 toxicity grading) worsened from Baseline will be summarized. The worst (maximum) CFB in toxicity grade for selected laboratory parameters will be presented in shift tables: the participant's baseline grade will be cross-tabulated by the participant's maximum postbaseline grade during the treatment, including scheduled and unscheduled visits.

5.6.3.3. Evaluation of Drug-Induced Serious Hepatotoxicity

A potential Hy's law case refers to ALT or AST > 3× the reference ULN, with bilirubin > 2× ULN. Possibly Hy's law cases can be visualized with use of evaluation of drug-induced serious hepatotoxicity plots, a log-log scatter plot where the x-axis is the peak postbaseline ALT as a multiple of ULN and the y-axis is the peak postbaseline bilirubin as a multiple of ULN. A drug-induced serious hepatotoxicity plot by quadrant will be presented by study intervention.

Summaries of liver enzymes will also be provided in a separate table. The number and percentage of participants in the following categories by visit, by any post-baseline and by worst post-baseline will be presented.

- ALT > 3×ULN, > 5×ULN, > 10×ULN or 20×ULN
- AST > 3×ULN, > 5×ULN, > 10×ULN or 20×ULN
- GGT > 3×ULN, > 5×ULN, > 10×ULN or 20×ULN
- ALT > 3×ULN + bilirubin > 2×ULN
- ALT > 3×ULN + ALP < 2×ULN + bilirubin >= 2×ULN

5.6.3.4. Physical Examination

Weight will be measured at each visit. Height was measured only once at Screening, and the same value will be used through the rest of the study accordingly.

5.6.3.5. Twelve-Lead ECGs

ECGs will be performed on a calibrated 12-lead machine according to the SoA. The following quantitative ECG measurements will be taken:

- ECG mean HR (beats per minute)
- PR interval (milliseconds)
- QRS duration (milliseconds)
- QT interval (milliseconds)
- QT corrected for heart rate by Bazett formula interval (milliseconds)
- QT corrected for heart rate by Fridericia's formula (QTcF) interval (milliseconds)
- RR interval (milliseconds)

The observed ECG measurements and CFB will be summarized descriptively by study intervention group and by visit. ECG results will also be classified as normal and abnormal and will be summarized. Shift tables from Baseline to each postbaseline visit will be generated by study intervention group using frequency tabulations. ECG results for QTcF will also be classified as ≤ 450 ms, > 450 to ≤ 480 ms, > 480 to ≤ 500 ms, or > 500 ms and will be summarized using frequency by study intervention and visit. As there are many participants with artificial pacemakers that affect the rate and intervals on the ECG, ECG data will be summarized by native rhythms separately.

A listing of ECG results will be presented by study intervention group, participant, and visit. Whether or not a participant has a pacer will be indicated in the listing.

5.7. Other Analyses

5.7.1. Additional Analyses for Secondary Endpoints

In order to align with the duration of the PETP (LPI+18M), the main analysis based on Ranked ANCOVA (Section 5.4.2.3) will be repeated for CFB to Week 76 for KCCQ-OS, GLS%, 6MWT, and SF-36v2 PCS. For NT-proBNP, which is measured every 4 weeks after Week 50 as per the SoA, the analysis will be repeated for CFB to Week 74.

Additionally, the sensitivity analysis using ANCOVA based on imputed data (Section 5.4.4) will also be repeated for CFB to Week 76 for KCCQ-OS, GLS%, 6MWT, SF-36v2 PCS. For NT-proBNP, which is measured every 4 weeks after Week 50 as per the SoA, the analysis will be repeated for CFB to Week 74.

5.7.2. Hematologic Response and Daratumumab Usage on Efficacy Endpoints

5.7.2.1. Hematologic Response

The assessment of the hematologic response profile is based on the following definitions (Palladini, 2021):

  - Complete response (CR) (both criteria must be met): Absence of amyloidogenic light chains (either free and/or as part of a complete immunoglobulin) defined by negative immunofixation electrophoresis of both serum and urine.
  - Either an FLC ratio within the reference range or the uninvolved FLC concentration is greater than the involved FLC concentration with or without an abnormal FLC ratio.
- Very good partial response (VGPR): dFLC < 4 mg/dL
- Partial response (PR): a dFLC decrease > 50% from pretreatment
- No response (NR): all others

Confirmation of complete hematologic response requires interpretation of serum and urine immunofixation results, which may potentially be confounded by concomitant administration of monoclonal antibodies of IgG kappa isotype, such as daratumumab or anselamimab. Therefore, additional analyses of deep hematologic response will be performed based on levels of dFLC and iFLC, independent of immunofixation results as follows:

- Proportion of participants with deep hematologic response (iFLC < 2 mg/dL or dFLC < 1 mg/dL, Wechalekar AD, 2023), VGPR, PR and NR will be provided by study intervention group at Weeks 4, 14, 26, 50. The difference in proportion of participants between the 2 intervention groups, p-value as well as the 2-sided 95% CI will be estimated using the CMH method adjusted by stratification variables: study, geographic region, and daratumumab usage at Baseline. No multiple imputation will be done for the missing data. If certain strata are empty or extremely sparse, clinically justifiable collapsing of those strata will be considered. Any such modifications to achieve adequate counts will be documented. The same analysis will be repeated based on the following categories: VGPR or better (dFLC < 4 mg/dL) and PR or NR (dFLC ≥ 4 mg/dL).
- Time to ACM (Section 5.4) will be assessed via the KM analysis by study intervention group and by deep hematologic response, VGPR, PR and NR. Participants will be categorized based on the best hematologic response achieved within PETP.
- Time to deep hematologic response, time to VGPR or better, time to PR or better will be assessed via the KM analysis. Patients without a hematologic response will be censored at the last FLC assessment date.

5.7.2.2. Daratumumab Usage

KM plots for time to ACM will also be provided by study intervention and daratumumab usage (yes or no). An analysis using the Cox proportional hazard model with time-dependent daratumumab usage may also be performed.

An analysis of the primary efficacy endpoint and key secondary efficacy endpoints will be performed for the following post-randomization subgroup.

- Use of daratumumab (versus no use of daratumumab) at any time during the study (study days ≥ -28)

5.7.3. Restricted Mean Survival Time

Restricted Mean Survival Time (RMST)-based method may be used to better understand the degree and nature of the treatment differences in terms of time to ACM between the 2 study intervention groups. RMST is defined as the area under the Kaplan-Meier survival curve up to Week 50. A stratified RMST analysis will be conducted to compare the treatment groups, using study, geographic region, and daratumumab usage at Baseline as stratification factors. The overall difference and SE in RMST between treatment groups will be presented. The corresponding p-value based on a stratified chi-square test will be presented.

5.7.4. Time to Composite Event of ACM or First CVH

Hazard ratio and its 95% CI of CAEL-101 versus placebo based on a Cox proportional hazard model stratified by study, geographic region and adjusted by daratumumab usage at Baseline will be reported for time to the composite event of ACM or First CVH. Time to the composite event of ACM or first CVH is defined as the weeks from the randomization date to the first CVH date or death date, whichever occurs first. Participants who are alive and without a CVH within PETP will be censored at the last known alive date. A KM plot of time to the composite event of ACM or first CVH, and p-value using the log-rank test, will also be provided. Participants with heart transplant or LVAD implantation will be treated as having died.

5.7.5. Time to First CVH

Time to first CVH is defined as the weeks from the randomization date to the first CVH date. Participants who are without any CVHs within PETP will be censored at the last known alive date. Hazard ratio and its 95% CI of CAEL-101 versus placebo based on a Cox proportional hazard model stratified by study, geographic region and adjusted by daratumumab usage at Baseline as well as a KM plot with p-value using the log-rank test will be generated.

5.7.6. Additional Cardiac Data Analysis

Descriptive statistics of the observed values and CFB over time may be reported for additional cardiac structure and function parameters, for example, left ventricular mass, interventricular septal thickness, and left ventricular ejection fraction. Cardiac magnetic resonance imaging data (only collected in Study CAEL101-302) will be analyzed: the CFB for Cardiac Extra-cellular Volume (%) at Baseline, Week 50, and last observation will be summarized in a table and a plot

of observed values as well as CFB will be generated. Participants with atrial fibrillation (and/or other arrhythmias) may also be summarized.

5.7.7. Subgroup Analyses

Analysis of the effects of CAEL-101 on various subgroups will be performed to assess whether there are any intrinsic or extrinsic characteristics that demonstrate clinically meaningful differences or trends. Analyses of the primary efficacy endpoint and key secondary efficacy endpoints will be provided for the following subgroups. SF-36v2 PCS and NT-proBNP will also be analyzed by subgroup of study. In addition, safety summaries will be provided for relevant subgroups (Section 5.6.2).

- Sex (male versus female).
- Race (American Indian or Alaska Native, Asian, Black or African American, Hawaiian or Other Pacific Islander, White or Caucasian, and Other).
- Age Group (below median age and above or equal to median age).
- Geographic region (North America versus the rest of the world): the analysis model should exclude geographic region as a covariate, if applicable.
- FLC subtype: Kappa or Lambda, whichever has a higher value at Baseline.
- Use of daratumumab (versus no use of daratumumab) as part of the first-line background anti-PCD treatment (-28 ≤ study day ≤ 1): the analysis model should remove use of daratumumab at Baseline, if applicable.
- GLS% value at Baseline (≤ -16.2%, > -16.2% to ≤ -12.2%, > -12.2% to < -9.0%, and ≥ -9.0%).
- NYHA Functional Classification (see CAEL101-301 Protocol Section 8.10 and CAEL101-302 Protocol Section 8.11) at Baseline.
- Study (CAEL101-301 and CAEL101-302): the analysis model should exclude study as a covariate, if applicable.

The pooled data are not powered to assess treatment outcomes by these subgroups, and thus, inferential assessments for the primary and secondary efficacy endpoints by subgroup should be interpreted with caution. Rather, trends across the subgroups will be reviewed for any similarities and/or differences.

A minimum of 15 participants per subgroup would be needed to perform the subgroup analyses. The subgroup analyses will be displayed using forest plots.

5.8. Interim Analyses

No interim analysis will be performed.

5.8.1. Independent Data Monitoring Committee

An IDMC will be established to monitor and review clinical safety data in an unblinded manner on a regular basis to ensure the continuing safety of the participants enrolled in these studies. The

IDMC will meet periodically to review safety data. Further details are available in the IDMC charter. No additional interim analysis is planned for efficacy or futility in these studies.

6. SUPPORTING DOCUMENTATION

6.1. Appendix 1: Technical Specifications for Derived Variables

6.1.1. Analysis Window

The analysis windowing for PK samples is provided in Table 5. Any unscheduled visits will be excluded from PK analysis (eg, tables and figures) but will be included in the data listings.

Table 5: Windowing for PK Samples

PK Samples		Window (Day)	
Visit	**Scheduled Study Day**	**Lower Limit**	**Upper Limit**
Baseline	1	1	1
Week 2	8	2	12
Week 3	15	13	19
Week 4	22	20	29
Week 6	36	30	50
Week > 6 (x = 10, 14, 18,…, every 4 weeks)	[(x-1) × 7] + 1	[(x - 1) × 7] - 12	[(x - 1) × 7] + 15
End of Treatment + Day 14	EOT + 14	EOT + 1	EOT + 21

For secondary and exploratory endpoints, Table 6, Table 7, Table 8, and Table 9 provide the windowing conventions for scheduled and unscheduled visits.

Note: since screening assessment results may be obtained within 28 days prior to randomization and the first dose of CAEL-101 must be administered within 24 hours after randomization, then the results may be considered within 29 days of the first dose of CAEL-101. If the window goes beyond 29 days, then the participant's screening data will be assessed by the clinical team to determine if there is any potential important protocol deviation that may affect the efficacy endpoints.

Table 6: Windowing for 6MWT, KCCQ, SF-36v2, and EQ-5D-5L

6MWT, KCCQ, SF-36v2, and EQ-5D-5L		Window (Day)	
Visit	**Scheduled Study Day**	**Lower Limit**	**Upper Limit**
Baseline[a]	-28 to 1	-28	1
Week 14	92	2	134
Week 26	176	135	218
Week 38	260	219	302
Week 50	344	303	435
Week > 50 (x = 76, 102, 128,…, every 26 weeks)	[(x-1) × 7] + 1	[(x - 1) × 7] - 89	[(x - 1) × 7] + 92
End of Treatment + Day 14	EOT + 14	EOT + 1	EOT + 21

[a] 6MWT is collected at Screening and not collected at Week 1. KCCQ, SF-36v2, and EQ-5D-5L are not collected at Screening and are collected at Week 1.

Abbreviations: 6MWT = 6-Minute Walk Test; EQ-5D-5L = European Quality of Life Health 5-Dimension 5-Level health questionnaire; KCCQ = Kansas City Cardiomyopathy Questionnaire; SF-36v2 = Short Form 36 Health Survey

Table 7: Windowing for Echocardiogram (Including GLS%), CRP, 24-Hour Urine Protein and NYHA Functional Classification

Echocardiogram, CRP, 24-Hour Urine Protein, and NYHA Functional Classification		Window (Day)	
Visit	**Scheduled Study Day**	**Lower Limit**	**Upper Limit**
Baseline	-28 to 1	-28	1
Week 14	92	2	134
Week 26	176	135	260
Week 50	344	261	435
Week > 50 (x = 76, 102, 128,…, every 26 weeks)	[(x-1) × 7] + 1	[(x - 1) × 7] - 89	[(x - 1) × 7] + 92
End of Treatment + Day 14[a]	EOT + 14	EOT + 1	EOT + 21
End of Treatment + Day 28	EOT + 28	EOT + 22	EOT + 42
End of Treatment + Day 56	EOT + 56	EOT + 43	EOT + 70
End of Treatment + Day 84	EOT + 84	EOT + 71	EOT + 98
End of Treatment + Day 112	EOT + 112	EOT + 99	EOT + 126
End of Treatment + Day 140	EOT + 140	EOT + 127	EOT + 154

[a] EOT echocardiogram (including GLS%), 24-hour urine protein, and NYHA functional classification are required ONLY at the 14 days following the last dose of study drug.
Abbreviations: CRP = C-reactive protein; GLS% = global longitudinal strain percentage; NYHA = New York Heart Association

Table 8: NT-proBNP, cTnT, and Serum FLC

NT-proBNP, cTnT, and Serum FLC		Window (Day)	
Visit	**Scheduled Study Day**	**Lower Limit**	**Upper Limit**
Baseline	-28 to 1	-28	1
Week 2	[(x-1) × 7] + 1	2	12
Week 3		13	19
Week 4		20	29
Week 6		30	50
Week x (x = 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62,…, every 4 weeks)		[(x-1) × 7] -12	[(x-1) ×7] +15
End of Treatment + Day 14	EOT + 14	EOT + 1	EOT + 21
End of Treatment + Day 28	EOT + 28	EOT + 22	EOT + 42
End of Treatment + Day 56	EOT + 56	EOT + 43	EOT + 70
End of Treatment + Day 84	EOT + 84	EOT + 71	EOT + 98
End of Treatment + Day 112	EOT + 112	EOT + 99	EOT + 126
End of Treatment + Day 140	EOT + 140	EOT + 127	EOT + 154

Abbreviations: cTnT = cardiac troponin T; NT-proBNP = N-terminal pro-B-type natriuretic peptide; FLC = free light chain

Table 9: Immunogenicity Test

Immunogenicity Test		Window (Day)	
Visit	**Scheduled Study Day**	**Lower Limit**	**Upper Limit**
Baseline	-28 to 1	-28	1
Week 10	[(x-1) × 7] + 1	2	120
Week 26		121	260
Week 50		261	435
Week > 50 (x = 76, 102, 128,…, every 26 weeks)		[(x - 1) × 7] - 89	[(x - 1) × 7] + 92
End of Treatment + Day 14	EOT + 14	EOT + 1	EOT + 21
End of Treatment + Day 28	EOT + 28	EOT + 22	EOT + 42
End of Treatment + Day 56	EOT + 56	EOT + 43	EOT + 70
End of Treatment + Day 84	EOT + 84	EOT + 71	EOT + 98
End of Treatment + Day 112	EOT + 112	EOT + 99	EOT + 126
End of Treatment + Day 140	EOT + 140	EOT + 127	EOT + 154

For the safety analysis, Table 10, Table 11, and Table 12 provide the windowing conventions for scheduled and unscheduled visits for clinical laboratory tests, 12-lead ECG, and vital signs, respectively.

Table 10: Windowing for Safety Clinical Laboratory Parameters

Safety Laboratory Tests		Window (Day)	
Visit	**Scheduled Study Day**	**Lower Limit**	**Upper Limit**
Baseline	-28 to 1	-28	1
Week 3	15	2	26
Week 6	36	27	50
Week x (x = 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62,…, every 4 weeks)	[(x - 1) × 7] + 1	[(x - 1) × 7] - 12	[(x - 1) × 7] + 15
End of Treatment + Day 14	EOT + 14	EOT + 1	EOT + 21
End of Treatment + Day 28	EOT + 28	EOT + 22	EOT + 42
End of Treatment + Day 56	EOT + 56	EOT + 43	EOT + 70
End of Treatment + Day 84	EOT + 84	EOT + 71	EOT + 98
End of Treatment + Day 112	EOT + 112	EOT + 99	EOT + 126
End of Treatment + Day 140	EOT + 140	EOT + 127	EOT + 154

Table 11: Windowing for 12-Lead ECG Parameters

12-Lead ECG		Window (Day)	
Visit	**Scheduled Study Day**	**Lower Limit**	**Upper Limit**
Baseline	-28 to 1	-28	1
Week 2	8	2	12
Week 3	15	13	19

12-Lead ECG		Window (Day)	
Visit	**Scheduled Study Day**	**Lower Limit**	**Upper Limit**
Week 4	22	20	99
Week 26	176	100	358
End of Treatment + Day 14	EOT + 14	EOT + 1	EOT + 21

Abbreviations: ECG = electrocardiogram

Table 12: Windowing for Vital Sign Parameters

Vital Signs		Window (Day)	
Visit	**Scheduled Study Day**	**Lower Limit**	**Upper Limit**
Baseline	-28 to 1	-28	1
Week 2	8	2	12
Week 3	15	13	19
Week 4	22	20	29
Week x (x = 6, 8,…,52)	[(x - 1) × 7] + 1	[(x - 1) × 7] - 5	[(x-1) ×7] + 8
End of Treatment + Day 14	EOT + 14	EOT + 1	EOT + 21
End of Treatment + Day 28	EOT + 28	EOT + 22	EOT + 42
End of Treatment + Day 56	EOT + 56	EOT + 43	EOT + 70
End of Treatment + Day 84	EOT + 84	EOT + 71	EOT + 98
End of Treatment + Day 112	EOT + 112	EOT + 99	EOT + 126
End of Treatment + Day 140	EOT + 140	EOT + 127	EOT + 154

In Table 13, the windowing convention is provided for the magnetic resonance imaging (MRI) data (only collected in Study CAEL101-302).

Table 13: Windowing for MRI

MRI		Window (Day)	
Visit	**Scheduled Study Day**	**Lower Limit**	**Upper Limit**
Baseline	-28 to 1	-28	1
Week 26	176	2	260
Week 50	344	261	526
Week 102	708	527	890
Week x (x = 154, …every 52 weeks)	[(x - 1) × 7] + 1	[(x - 1) × 7] - 180	[(x-1) ×7] + 183
End of Treatment + Day 14	EOT + 14	EOT + 1	EOT + 21

Abbreviations: MRI = magnetic resonance imaging

6.1.2. Selection of Multiple Records in an Analysis Window

If multiple valid, nonmissing measurements exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- In general, baseline is defined as the last nonmissing assessment for a given parameter prior to the first study intervention infusion.
    - If the assessment time is available for a given parameter, the last nonmissing value prior to the first dose date-time will be used as baseline.
    - If the assessment time is not available for a given parameter:
        - If the screening assessments are not missing, then the last screening measurement value prior to the first dose date will be used as baseline.
        - If the screening assessments are missing and the first assessment is collected on the first dose date, then the measurement value on the first dose date will be used as baseline.
    - If multiple measurements occur on the same day, the last nonmissing value prior to the time of the first dose of study intervention will be considered as the baseline value. If these multiple measurements occur at the same time or the time is not available, the average of these measurements (for continuous data) will be considered as the baseline value.
- For postbaseline visits:
    - The record closest to the scheduled day for that visit will be selected.
    - If there are 2 records that are equidistant from the scheduled day, the later record will be selected.
    - If there is more than 1 record on the selected day, the later in time will be selected, unless otherwise specified.

6.1.3. Adverse Event

AEs occurring on or after the start of the first infusion and up to the last infusion date + 140 days will be referred to as TEAEs and will be the focus of the AE analysis. Outcomes will be reported for the SS.

- TEAEs are events with start dates and start times on or after the date and time of the start of the first study intervention dose. If the start date of an AE is partially or completely missing and the end date and time of the AE does not indicate that it occurred prior to the first dose, then the determination of treatment-emergent status will be based on the following:
    - If the start year is after the year of the first study intervention dose, then the AE is treatment emergent; else
    - If the start year is the same as the year of the first study intervention dose and the start month is missing, then the AE is treatment emergent; else
    - If the start month is present and is the same or after the month of the first study intervention dose, then the AE is treatment emergent; else
    - If the start date is completely missing, then the AE is treatment emergent.
- All other AEs are considered pretreatment AEs.

To be able to calculate the time from the first dose to an AE, the following are the imputation rules for AE start dates:

Imputation Rules for Partial Dates			
Parameter	**Missing**	**Additional Conditions**	**Imputation**
Start date for AEs	D	M and Y are the same as the M and Y of the first dose of study intervention.	Date of the first dose of study intervention
		M and/or Y are not the same as the date of the first dose of study intervention.	First day of the month
	D and M	Y is the same as the Y of the first dose of study intervention.	Date of the first dose of study intervention
		Y is after the Y of the first dose.	Set to 01 Jan
	D, M, and Y	None - Date is completely missing.	Date of the first dose of study intervention

D = day; M = month; Y = year

AE duration (days) = date of stop of AE – date of start of AE + 1

Duration will be set to "missing" if the stop date of an AE is incomplete or if the AE is ongoing.

6.1.4. Imputation of Missing Date of Death

In the event of a participant's death, the complete date of death, including the day, month, and year, will be obtained. If the date of the death is missing or partially missing, the following imputation rules should be applied:

Imputation Rules for Death Dates		
Missing	**Additional Conditions**	**Imputation**
D	M and Y are the same as the M and Y of the last known date recorded.	Date of the last known date recorded + 1 day
	M and/or Y is not the same as the last known date recorded.	First day of the month
D and M	Y is the same as the Y of the last known date recorded.	Date of the last known date recorded +1 day
	Y is not the same as the last known date recorded.	Set to 01 Jan
D, M, and Y	None – date is completely missing	Last known date recorded +1 day

D = day; M = month; Y = year

6.1.5. Imputation of Missing Dates of CVH

If the admission date of the CVH is missing or partially missing, the following imputation rules will be applied:

Imputation Rules for CVH Admission Date		
Missing	**Additional Conditions**	**Imputation**
D	M and Y are the same as the M and Y of the corresponding AE start date.	Date of the corresponding AE start date
	M and/or Y is after the corresponding AE start date.	First date of the month
D and M	Y is the same as the Y of the corresponding AE start date.	Date of the corresponding AE start date
	Y is after the Y of the corresponding AE start date.	Set to 01 Jan
D, M, and Y	None – Date is completely missing.	Date of the corresponding AE start date

6.1.6. Estimated Glomerular Filtration Rate

The derivation of eGFR will be based on the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) Creatine Equation (2021) as follows:

$$eGFR_{cr} = 142 \times \min(S_{cr}/\kappa, 1)^{\alpha} \times \max(S_{cr}/\kappa, 1)^{-1.200} \times 0.9938^{Age} \times 1.012 \text{ [if female]}$$

where

S_{cr} = standardized serum creatinine in mg/dL

κ = 0.7 (females) or 0.9 (males)

α = -0.241 (female) or -0.302 (male)

$\min(S_{cr}/\kappa, 1)$ is the minimum of S_{cr}/κ or 1.0

$\max(S_{cr}/\kappa, 1)$ is the maximum of S_{cr}/κ or 1.0

Age (years)

eGFR categories in chronic kidney disease will be assigned based on eGFR values as in the table below:

eGFR Category	**eGFR (mL/min/1.73 m^2)**	**Terms**
G1	≥ 90	Normal or high
G2	60 to < 90	Mildly decreased
G3a	45 to < 60	Mildly to moderately decreased
G3b	30 to < 45	Moderately to severely decreased
G4	15 to < 30	Severely decreased
G5	< 15	Kidney failure

6.2. Appendix 2: Study and Participant Characteristics

6.2.1. Protocol Deviations

Protocol deviations will be presented in a listing that includes the severity, protocol deviation category, and action taken, if any. Any protocol deviation determined to have an impact on participant safety, ICH-Good Clinical Practice compliance, or efficacy endpoints may be categorized as an important protocol deviation per the study team's approval prior to the database lock during the blinded data review meeting. The decision whether a protocol deviation is an important deviation or not will be made on a case-by-case basis.

Only important protocol deviations that may affect the interpretation of study outcomes will be used to exclude participants from the PP Analysis Set. For instance, if the informed consent form is signed after some screening study procedures are performed, this may (or may not) be an important protocol violation. If it is assigned as an important violation, this would not, however, affect the interpretability of the efficacy outcomes and thus would not result in exclusion of the participant from the PP Analysis Set.

The possible reasons of important protocol deviations are listed but not limited to the following:

- Failure to obtain informed consent
- Deviation from inclusion/exclusion criteria
- Intake of prohibited medication (refer to Protocol Section 6.9 for a list of prohibited medications). Only participants who have used doxycycline for more than 4 weeks will be excluded from the PP Analysis Set.
- Missing more than 20% of the total planned study drug doses (eg, a participant missed 25% of the total planned study drug doses during the study accumulatively)
- Participants receiving the wrong medication kit or incorrect treatment at any visit
- Protocol waivers that may affect the primary or secondary endpoint analyses
- Unintended breaking of the blind: any member of the study team and/or any individual at the investigator site is unintentionally unblinded

Important protocol deviations will be summarized by study intervention group based on the ITT Analysis Set. All participants with protocol deviations, including those specified above, will be listed.

The number and percentage of participants with specific important protocol deviations resulting in exclusion from the PP Analysis Set will be summarized by study intervention group and overall using the ITT Analysis Set.

Protocol deviations from monitoring reports and other relevant sources will also be reviewed, and any important deviations will be included in the list that is summarized and reported.

6.2.2. Demographics, Disease Characteristics, and History

All demographic, baseline characteristics information, and disease (eg, AL amyloidosis) characteristics at Baseline will be summarized descriptively (eg, the number and percentage of participants for each category for categorical parameters and the number of participants with

nonmissing values, mean, SD, lower quartile, median, upper quartile, IQR, and range for continuous parameters) using the SS and ITT Analysis Sets. No formal statistical testing will be performed to compare differences between study intervention groups.

Summary statistics will be presented by study intervention group and overall. It will also be reported by study.

6.2.2.1. Demographics and Baseline Characteristics

The following demographic variables, collected at Baseline, will be summarized by study intervention group and overall using descriptive statistics:

- Sex
- Race
- Ethnicity
- Age (years)
- Age (using the study median age)
- Age group (< 65 years versus ≥ 65 years)
- Age group (< 65 years versus 65 to < 75 years versus ≥ 75 years)
- Geographic region (North America and the rest of the world)

The following baseline characteristics variables will be summarized by intervention group and overall using descriptive statistics:

- Weight
- Height
- Body mass index
- BSA

BSA is calculated using Mosteller's formula as follows:

$$BSA\ (m^2) = \sqrt{\frac{height\ (cm) \times weight\ (kg)}{3600}}$$

6.2.2.2. Disease Characteristics

The following AL amyloidosis characteristics at Baseline will be summarized descriptively:

- Mayo stage (IIIa and IIIb)
- NT-proBNP
- cTnT
- FLC subtype (kappa and lambda)
- FLC ratio

- dFLC
- KCCQ-OS
- SF-36v2 PCS Norm-based T score
- 6MWT
- GLS%
- Renal stage, defined by the following criteria, in participants who have renal involvement (ie, baseline urine protein > 0.5 g/24 h).
  - Renal Stage I: eGFR ≥ 50 mL/min and proteinuria ≤ 5 g/24 h
  - Renal Stage II: eGFR < 50 mL/min or proteinuria > 5 g/24 h
  - Renal Stage III: eGFR < 50 mL/min and proteinuria > 5 g/24 h

The Mayo stage is based on the NT-proBNP and cTnT results obtained at the time of Screening as described in Table 14. It is expected that participants will be classified according to their study assignment (ie, Study 302 [for stage IIIa participants]). A summary of the Mayo staging and corresponding parameters at Screening will be generated. Any participants who are misclassified will be included in a by-participant data listing with these staging parameters listed.

Table 14: Staging of AL Amyloidosis with Advanced Cardiac Involvement for CARES Studies Based on the European Modification of the 2004 Mayo Staging (Palladini, 2016, Adapted from Muchtar, 2019)

Stage I	Stage II	Stage IIIa	Stage IIIb
Zero markers above the threshold: • NT-proBNP < 332 ng/L AND one of the following: • hs-cTnT < 50 pg/mL (0.05 ng/mL) • cTnT ≤ 0.035 ng/mL • cTnI ≤ 0.1 ng/mL	One marker above the threshold: • NT-proBNP ≥ 332 ng/L OR one of the following: • hs-cTnT ≥ 50 pg/mL (0.05 ng/mL) • cTnT ≥ 0.035 ng/mL • cTnI ≥ 0.1 ng/mL	Two markers above the threshold: • 332 ng/L ≤ NT-proBNP[a] ≤ 8500 ng/L AND one of the following: • hs-cTnT ≥ 50 pg/mL (0.05 ng/mL) • cTnT ≥ 0.035 ng/mL • cTnI ≥ 0.1 ng/mL	Two markers above the threshold: • NT-proBNP > 8500 ng/L AND one of the following: • hs-cTnT ≥ 50 pg/mL (0.05 ng/mL) • cTnT ≥ 0.035 ng/mL • cTnI ≥ 0.1 ng/mL

Note: The 2013 European Modification of the 2004 Standard Mayo Clinic Staging in participants with advanced cardiac involvement is based on the conventional (generation 4) cTnT assay. However, hs-cTnT, or the generation 5 assay, is becoming more widely available and is used more commonly in clinical practice. It is established that a cTnT value of > 0.035 ng/mL can be extrapolated to an hs-cTnT value of ≥ 50 pg/mL (0.05 ng/mL) for the determination of Mayo stage (Muchtar, 2019).

[a] For the CAEL101-302 study, the threshold for NT-proBNP is ≥ 650 and ≤ 8500 ng/L

Abbreviations: AL = amyloid light chain; CARES = Cardiac Amyloid Reaching for Extended Survival; cTnI = cardiac troponin I; cTnT = cardiac troponin T; hs-cTnT = high-sensitivity cardiac troponin T; NT-proBNP = N-terminal pro-B-type natriuretic peptide

6.2.2.3. Medical History

Relevant and significant medical history and concurrent illness will be collected for all participants and noted as to whether the condition is active at the first visit of the Screening Period. Medical history information will be coded to the primary SOC and PT using the MedDRA (Version 23.0 or higher). The number and percentage of participants with medical history events will be summarized by body system and PT. The frequency and percentage of medical history and concomitant medical conditions will be summarized by SOC and PT for the ITT and SS Analysis Sets. The number (%) of participants reporting each condition will be presented by overall and study intervention group. If participants have more than 1 disease within an SOC or PT, they will be counted only once for the respective SOC or PT.

To be able to calculate the amount of time since the diagnosis of a medical condition, the imputation rules for the missing start date of medical history are as follows:

Imputation Rules for Partial Dates (D = day, M = month, Y = year)		
Parameter	**Missing**	**Imputation**
Start date for medical history	D	First day of the month
	D and M	Set to 01 Jan
	D, M, and Y If the medical condition is ongoing If the medical condition is not ongoing	 Informed consent date No imputation

6.2.3. Prior and Concomitant Medications/Therapies

Medications will be mapped to a generic term using the World Health Organization Drug Dictionary of 15 Mar 2018 or later. A prior medication is defined as a medication with an end date before the first dose of study intervention. According to the protocol, medications up to 30 days prior to informed consent will be collected. All other medications will be considered concomitant. Concomitant medications will include medications that are ongoing or with insufficient date data (partial/missing dates) to determine that it was a prior medication. Note that PCD medications will be analyzed separately from other medications.

The World Health Organization Drug Dictionary will be used to classify medications by PT, and World Health Organization Anatomical Therapeutic Chemical (ATC) classification of ingredients. ATC Level 3 will be used.

The following steps will be implemented to categorize the medications:

Algorithm for the Categorization of Medications (Prior or Concomitant)			
Parameter	**Value**	**Additional Conditions**	**Medication Category/Action**
Ongoing flag	Yes	Not applicable.	Concomitant
	No	The medication end date is partial or missing.	Perform the imputation on the medication end date and then assign the medication category
		The medication end date is before the first dose of study intervention.	Prior

Algorithm for the Categorization of Medications (Prior or Concomitant)			
Parameter	**Value**	**Additional Conditions**	**Medication Category/Action**
		The medication end date is on or after the first dose of study intervention.	Concomitant

The following imputation rules for the end date of the medication will be applied to medications other than PCD medications:

Imputation Rules for Partial Dates (D = day, M = month, Y = year)			
Parameter	**Missing**	**Additional Conditions**	**Imputation**
End date of medication	D	M and Y are the same as the M and Y of the first dose of study intervention If the start date of medication is not missing If the start date of medication is partially missing or completely missing	 Day of the later of (date of the first dose of study intervention, start date of medication) Date of the first dose of study intervention
		M and/or Y is not the same as the date of the first dose of study intervention	Last day of the month
	D and M	Y is the same as the Y of the first dose of study intervention If the start date of medication is not missing If only the D of the start date of medication is missing If the D and M of the start date of medication is missing or the start date is completely missing	 Day and month of the later of (date of the first dose of study intervention, start date of medication) First day and month of the start date of medication Date of the first dose of study intervention
		Y is before/after the Y of the first dose If the start date of medication is not missing If only the D of the start date of medication is missing If the D and M of the start date of medication are missing or completely missing	 Day and month of the start date of medication First day and month of the start date of medication 01 Jan
	D, M, and Y	None – Date is completely missing	Later of (date of the first dose of study intervention, start date of medication)

D = day; M = month; Y = year

The following imputation rules for the start date of medication will be applied to all medications:

Imputation Rules for Partial Dates (D = day, M = month, Y = year)			
Parameter	**Missing**	**Additional Conditions**	**Imputation**
Start date of medication	D	M and Y are the same as the M and Y of the first dose of study intervention If the end date of the medication is not missing If the end date of the medication is missing	 Day of the earlier of (date of the first dose of study intervention, end date of medication) Date of the first dose of study intervention
		M and/or Y is not the same as the date of the first dose of study intervention	First day of the month
	D and M	Y is earlier or the same as the Y of the first dose of study intervention If the end date of the medication is not missing If the end date of the medication is missing	 Day and month of the earlier of (date of the first dose of study intervention, end date of medication) Day and month of the date of the first dose of study intervention
		Y is after the Y of the first dose If the end date of medication is not missing If D and M of the end date of medication is missing	 Day and month of the end date of medication 01 Jan
	D, M, and Y	None - Date is completely missing If the end date of medication is not missing If the end date of medication is missing	 Earlier of (date of the first dose of study intervention, end date of medication) Date of the first dose of study intervention

D = day; M = month; Y = year

The following imputation rules for end date of medication will be applied to PCD medications:

Imputation Rules for Partial Dates (D = day, M = month, Y = year)			
Parameter	**Missing**	**Additional Conditions**	**Imputation**
End date of PCD medication	D	Same as other medications	
	D and M	Same as other medications	
	D, M, and Y	None - Date is completely missing If ongoing flag = 'Y' If the end of study date is missing If the end of study date is not missing	 End date of PETP End of study date

D = day; M = month; Y = year

Prior and concomitant medications will be summarized separately by drug class (ATC Level 3) and PT.

The listing will show all prior and concomitant medications and therapies used during the study.

Prior and concurrent PCD treatments will be summarized (separately) by duration and by study intervention group and overall. The following will be summarized:

- Duration of PCD therapy
- Number (%) of participants taking daratumumab by country (or region)
- Duration of daratumumab
- Hematologic response by daratumumab versus no daratumumab subgroups

6.3. Appendix 3: Instrument Scoring Details

6.3.1. KCCQ Scoring

The KCCQ scoring is provided below.

1. Physical Limitation

- Code responses to each of Questions 1a to 1f as follows:

  Extremely limited = 1
  Quite a bit limited = 2
  Moderately limited = 3
  Slightly limited = 4
  Not at all limited = 5
  Limited for other reasons or did not do = <*missing value*>

- If at least 3 of Questions 1a to 1f are not missing, then compute

  Physical Limitation Score = 100 × [(mean of Questions 1a-f actually answered) - 1] / 4

  (See the footnote at the end of this section for the explanation of "mean of questions actually answered.")

2. Symptom Stability

- Code the responses to Question 2 as follows:

  Much worse = 1
  Slightly worse = 2
  Not changed = 3
  Slightly better = 4
  Much better = 5
  I've had no symptoms over the last 2 weeks = 3

- If Question 2 is not missing, then compute

  Symptom Stability Score = 100 × [(Question 2) - 1] / 4

3. Symptom Frequency

- Code responses to Questions 3, 5, 7, and 9 as follows:

  Question 3
  Every morning = 1
  3 or more times a week but not every day = 2
  1-2 times a week = 3
  Less than once a week = 4
  Never over the past 2 weeks = 5

  Questions 5 and 7
  All of the time = 1
  Several times a day = 2
  At least once a day = 3
  3 or more times a week but not every day = 4
  1-2 times a week = 5
  Less than once a week = 6
  Never over the past 2 weeks = 7

  Question 9
  Every night = 1
  3 or more times a week but not every day = 2
  1-2 times a week = 3
  Less than once a week = 4
  Never over the past 2 weeks = 5

- If at least 2 of Questions 3, 5, 7, and 9 are not missing, then compute:

  S3 = [(Question 3) - 1]/4
  S5 = [(Question 5) - 1]/6
  S7 = [(Question 7) - 1]/6
  S9 = [(Question 9) - 1]/4

Symptom Frequency Score = 100 × (mean of S3, S5, S7, and S9)

4. Symptom Burden

- Code the responses to Questions 4, 6, and 8 as follows:

  Extremely bothersome = 1
  Quite a bit bothersome = 2
  Moderately bothersome = 3
  Not at all bothersome = 4
  I've had no swelling/fatigue/shortness of breath = 5

- If at least one of Questions 4, 6, and 8 is not missing, then compute

  Symptom Burden Score = 100 × [(mean of Questions 4, 6, and 8 actually answered) - 1] / 4

5. Total Symptom Score

= mean of the following available summary scores:

Symptom Frequency Score

Symptom Burden Score

6. Self-Efficacy

- Code responses to Questions 10 and 11 as follows:

  Question 10
  Not at all sure = 1
  Not very sure = 2
  Somewhat sure = 3
  Mostly sure = 4
  Completely sure = 5

  Questions 11
  Do not understand at all = 1
  Do not understand very well = 2
  Somewhat understand = 3
  Mostly understand = 4
  Completely understand = 5

- If at least one of Questions 10 and 11 is not missing, then compute

  Self-Efficacy Score = 100 × [(mean of Questions 10 and 11 actually answered) - 1] / 4

7. Quality of Life

- Code responses to Questions 12, 13 and 14 as follows:

  Question 12
  It has extremely limited my enjoyment of life = 1
  It has limited my enjoyment of life quite a bit = 2
  It has moderately limited my enjoyment of life = 3
  It has slightly limited my enjoyment of life = 4
  It has not limited my enjoyment of life at all = 5

Questions 13
Not at all satisfied = 1
Mostly dissatisfied = 2
Somewhat dissatisfied = 3
Mostly satisfied = 4
Completely satisfied = 5

Questions 14
I felt that way all of the time = 1
I felt that way most of the time = 2
I occasionally felt that way = 3
I rarely felt that way = 4
I never felt that way = 5

- If at least one of Questions 12, 13, and 14 is not missing, then compute

  QoL Score = 100 × [(mean of Questions 12, 13, and 14 actually answered) - 1] / 4

8. Social Limitation

- Code responses to each of Questions 15a to 15d as follows:

  Severely limited = 1
  Limited quite a bit = 2
  Moderately limited = 3
  Slightly limited = 4
  Did not limit at all = 5
  Does not apply or did not do for other reasons = <*missing value*>

- If at least 2 of Questions 15a to 16 are not missing, then compute

  Social Limitation Score = 100 × [(mean of Questions 15a to 16 actually answered) - 1] / 4

9. Overall Summary Score

= mean of the following available summary scores:

Physical Limitation Score
Total Symptom Score
QoL Score
Social Limitation Score

10. Clinical Summary Score

= mean of the following available summary scores:

Physical Limitation Score
Total Symptom Score

Note: references to "**means of questions actually answered**" imply the following.

- If there are *n* questions in a scale and the participant must answer *m* to score the scale but the participant answers only *n* - *i*, where *n* - *i* ≥ *m*, calculate the **mean of those questions** as
  (sum of the responses to those *n* - *i* questions) / (*n* - *i*)
  not
  (sum of the responses to those *n* - *i* questions) / n

6.3.2. European Quality of Life Health 5-Dimension 5-Level Questionnaire

EQ-5D-5L is a descriptive system assessing 5 dimensions: mobility, self-care, usual activity, pain/discomfort, anxiety/depression. Within each dimension, there are 5 response levels (see Table 15): no problems (Level 1), slight problems (Level 2), moderate problems (Level 3), severe problems (Level 4), and unable to perform activity or have extreme problems in the dimension (Level 5). The respondent is asked to indicate their health state by checking the box next to the most appropriate response level for each of the 5 dimensions. Responses are coded as single-digit numbers expressing the severity level selected in each dimension. For instance, "slight problems (level 2)" (eg, "I have slight problems in walking about") is always coded as "2." The digits for the 5 dimensions can be combined in a 5-digit code that describes the respondent's health state; for instance, 21111 means slight problems in the mobility dimension and no problems in any of the other dimensions. These numbers representing the severity levels of the 5 dimensions are labels used in the numerical description of a health state. They have no arithmetic properties. Therefore, these numbers should not be used to derive a summary score. Instead, an index value is introduced.

Index Score

After obtaining a 5-digit code that represents the 5 dimensions for a participant, an index value is calculated by applying a formula that attaches values (also called weights) to each of the levels in each dimension. The index value can then be calculated by deducting the appropriate weights of the 5 dimensions from 1. The corresponding weight used in the formula for each level within each dimension is illustrated in Table 15. For example, for a participant coded with health state 21354, the index value = 1 - 0.096 - 0 - 0.101 - 0.414 - 0.299 = 0.090; for a participant coded with health state 11111, the index value = 1 - 0 - 0 - 0 – 0 - 0 = 1.

Table 15: Index Score Calculation Using cTTO

US cTTO	Weights
Full health (11111)	
Mobility	
Level 1 (no problems)	0
Level 2 (slight problems)	-0.096
Level 3 (moderate problems)	-0.122
Level 4 (severe problems)	-0.237
Level 5 (unable to walk)	-0.322

US cTTO	Weights
Self-care	
Level 1 (no problems)	0
Level 2 (slight problems)	-0.089
Level 3 (moderate problems)	-0.107
Level 4 (severe problems)	-0.220
Level 5 (unable to self-care)	-0.261
Usual activity	
Level 1 (no problems)	0
Level 2 (slight problems)	-0.068
Level 3 (moderate problems)	-0.101
Level 4 (severe problems)	-0.255
Level 5 (unable to do usual activities)	-0.255
Pain/discomfort	
Level 1 (no pain or discomfort)	0
Level 2 (slight pain or discomfort)	-0.060
Level 3 (moderate pain or discomfort)	-0.098
Level 4 (severe pain or discomfort)	-0.318
Level 5 (extreme pain or discomfort)	-0.414
Anxiety/depression	
Level 1 (not anxious or depressed)	0
Level 2 (slightly anxious or depressed)	-0.057
Level 3 (moderately anxious or depressed)	-0.123
Level 4 (severely anxious and depressed)	-0.299
Level 5 (extremely anxious and depressed)	-0.321

Abbreviation: cTTO = composite time trade-off
Source: Pickard, 2019

Visual Analog Score

The European Quality visual analog score is a scale ranging from 0 to 100 where participants are asked to indicate their overall health on the day of questionnaire completion on a vertical visual analog score, anchored at each end by “the worst health you can imagine” (0) to “the best health you can imagine” (100).

6.4. Appendix 4: Additional Details on Statistical Methods

6.4.1. Intercurrent Events

Multiple ICE strategies are used in Section 5.3.2 and 5.4.1 for the primary and secondary estimands. Below are the descriptions in ICH E9 (R1) Addendum (ICH, 2021) for each ICE strategy.

Treatment Policy Strategy

The occurrence of the ICE is considered irrelevant in defining the treatment effect of interest; the values for the variable of interest are used, regardless of whether the ICE occurs. For example, when specifying how to address the use of additional medication as an ICE, the values of the variable of interest should be used, whether the participant takes additional medication or not.

If applied in relation to whether a participant continues treatment and whether a participant experiences changes in other treatments (eg, background or concomitant treatments), the ICE is considered to be part of the treatments being compared. In that case, this reflects the comparison described in the ICH E9 (ICH, 1998) Glossary (under the ITT principle) as the effect of a treatment policy.

In general, the treatment policy strategy cannot be implemented for ICEs that are terminal events, because values for the variable after the ICE do not exist. For example, an estimand based on this strategy cannot be constructed with respect to a variable that cannot be measured because of death.

Composite-Variable Strategies

This strategy relates to the variable of interest. An ICE is considered to be informative about the participant's outcome and is therefore incorporated into the definition of the variable. For example, a participant who discontinues treatment because of toxicity may be considered not to have been successfully treated. If the outcome variable was already a success or failure, discontinuation of treatment for toxicity would simply be considered another mode of failure. Composite-variable strategies do not need to be limited to dichotomous outcomes, however. For example, in a study measuring PF, a variable might be constructed using outcomes on a continuous scale, with participants who died being attributed a value reflecting the lack of ability to function. Composite-variable strategies can be viewed as implementing the ITT principle in some cases when the original measurement of the variable might not exist or might not be meaningful, but where the ICE itself meaningfully describes the participant's outcome, such as when the participant dies. Terminal events, such as death, are perhaps the most salient examples of the use of a composite strategy. If a treatment saves lives, its effect on various measures in surviving participants may be of interest, but it would be inappropriate to say that the summary measure of interest was only the average value of some numerical measure in survivors. The outcome of interest is survival along with the numerical measures. For example, progression-free survival in oncology studies measures the treatment effect on a combination of the growth of the tumor and survival.

While-on-Treatment Strategies

For this strategy, response to treatment before the occurrence of the ICE is of interest. Terminology for this strategy will depend on the ICE of interest (eg, while alive or when considering death as an ICE).

If a variable is measured repeatedly, its values up to the time of the ICE may be considered relevant for the clinical question, rather than the value at the same fixed timepoint for all participants. The same applies to the occurrence of a binary outcome of interest up to the time of the ICE. For example, participants with a terminal illness may discontinue a purely symptomatic treatment because they die, yet the success of the treatment can be measured based on the effect

on symptoms before death. Alternatively, participants might discontinue treatment, and in some circumstances, it will be of interest to assess the risk of an adverse drug reaction while the participant is exposed to treatment.

Similar to the composite-variable strategy, the while-on-treatment strategy can hence be thought of as having an impact on the definition of the variable (in this case, by restricting the observation time of interest to the time before the ICE). Particular care is important if the occurrence of the ICE differs between the treatments being compared.

6.4.2. Ranked ANCOVA

If higher values of a measurement indicate better outcomes (eg, KCCQ-OS), use the approach outlined below to obtain the values to use for ranking. Let $\hat{X}_{ij}$ denote the value to use in the ranked analysis for the j^{th} participant in the i^{th} treatment group (Lachin, 1999) at Week 50.

$$\hat{X}_{ij} = (1 - \delta_{ij})X_{ij} + \delta_{ij}(\eta + t_{ij})$$

where δ_{ij} = 1, where there is a missing value at Week 50, and δ_{ij}= 0, where there is a nonmissing value at Week 50. X_{ij} is the nonmissing value at Week 50. The table below lists the corresponding η and $t_{ij}$for the scenarios listed above for participants with missing values.

	η Value	**t_{ij}**
Participants who have missing values at Week 50 and died or had a heart transplant or LVAD implantation prior to Week 50	-99999	Time to death (weeks)
Participants who have missing values at Week 50 and vital status is unknown at Week 50	-88888	Time on study intervention (weeks)
Participants who have missing values at Week 50 and are known to be alive at Week 50	-77777	Time on study intervention (weeks)

If higher values of a measurement indicate worse outcomes (eg, GLS%), use the formula and table below to obtain the values for ranking. Note that the absolute value of η can be increased to ensure correct ranking in case of extreme observed values.

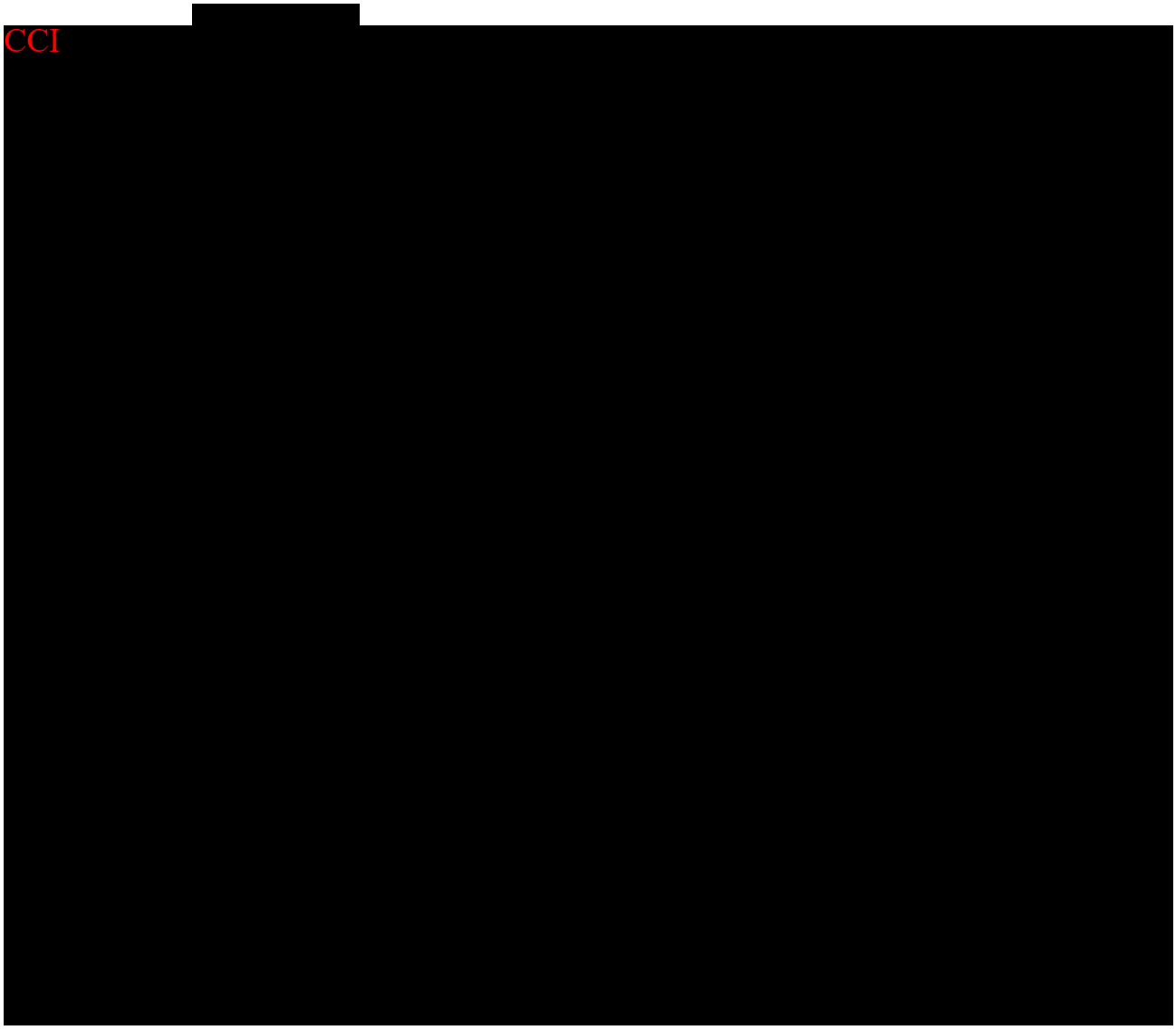
$$\hat{X}_{ij} = (1 - \delta_{ij})X_{ij} + \delta_{ij}(\eta - t_{ij})$$

	η Value	**t_{ij}**
Participants who have missing values at Week 50 and died or had a heart transplant or LVAD implantation prior to Week 50	99999	Time to death (weeks)
Participants who have missing values at Week 50 and vital status is unknown at Week 50	88888	Time on study intervention (weeks)
participants who have missing values at Week 50 and are known to be alive at Week 50	77777	Time on study intervention (weeks)

$\widehat{X_{ij}}$will be ranked from the worst outcome to the best outcome before being submitted to the Ranked ANCOVA model (Section 5.4.2.3).

6.4.3. Finkelstein-Schoenfeld Method

Table 16: Finkelstein-Schoenfeld Example Scoring Algorithm CCI



6.5. Appendix 5: Changes to Protocol-planned Analyses

Analysis of hepatic progression will not be done as defined in the protocol. The analysis may be done post hoc if deemed clinically relevant.

7. REFERENCES

Cox DR. Regression models and life-tables. J R Stat Soc. Series B (Methodological). 1972;34(2):187-220. JSTOR. Available from: www.jstor.org/stable/2985181. Accessed 20 Jul 2021.

Comenzo RL, Reece D, Palladini G, et al. Consensus guidelines for the conduct and reporting of clinical trials in systemic light-chain amyloidosis. Leukemia. 2012;26(11):2317-2325.

Dong G, Qiu J, Wang D, Vandemeulebroecke M. The stratified win ratio. J Biopharm Stat. 2018;28(4):778-796. doi: 10.1080/10543406.2017.1397007. Epub 2017 Nov 27. PMID: 29172988.

Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. Statist Med. 1999;18;1341-54.

Green P, Porter C, Bresnahan D, Spertus J. Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. J Am Coll Cardiol. 2000;35(5):1245-1255).

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). E9 Statistical Principles for Clinical Trials, September 1998. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e9-statistical-principles-clinical-trials. Accessed 21 Nov 2021.

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials, Guidance for Industry, May 2021. Available from: https://www.fda.gov/media/148473/download. Accessed 16 Dec 2021.

Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958;53(282):457-81. JSTOR. Available from: www.jstor.org/stable/2281868. Accessed 20 Jul 2021.

Lachin JM. Worst-rank score analysis with informatively missing observations in clinical trials. Control Clin Trials. 1999 Oct;20(5):408-22. doi: 10.1016/s0197-2456(99)00022-7. PMID: 10503801.

Maurer M, Schwartz J, Gundapaneni B, Elliott P, Merlini G, Waddington-Cruz M, et al. Tafamidis treatment for patients with transthyretin amyloid cardiomyopathy. N Engl J Med. 2018; 379:1007-1016.

Muchtar E, Dispenzieri A, Leung N, et al. Depth of organ response in AL amyloidosis is associated with improved survival: grading the organ response criteria. Leukemia. 2018;32(10):2240-2249.

Muchtar E, Kumar SK, Gertz MA, et al. Staging systems use for risk stratification of systemic amyloidosis in the era of high-sensitivity troponin T assay. Blood. 2019;133(7):763-766.

Muchtar E, Wisniowski B, Palladini G, et al. Graded renal response criteria for light chain (AL) amyloidosis. Blood. 2021;138 (Supplement 1):2721-2721.

Muchtar E, Dispenzieri A, Wisniowski B. Graded cardiac response criteria for patients with systemic light chain amyloidosis. J Clin Oncol. 2023;41(7):1393-1403.

Palladini G, Hegenbart U, Milani P, et al. A staging system for renal outcome and early markers of renal response to chemotherapy in AL amyloidosis. Blood. 2014;124(15):2325-2332.

Palladini G, Merlini G. What is new in diagnosis and management of light chain amyloidosis? Blood. 2016;128(2):159-168.

Palladini G, Schönland SO, Sanchorawala V, et al. Clarification on the definition of complete haematologic response in light-chain (AL) amyloidosis. Amyloid. 2021;28(1):1-2.

Pickard SA, Law EH, Jiang R, et al. United States valuation of EQ-5D-5L health states using an international protocol. Value Health. 2019;22:931-41.

Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. Eur Heart J. 2012 Jan;33(2):176-182.

Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons; 1987.

US Food and Drug Administration (FDA). Medical Device Development Tool (MDDT) Qualification Decision Summary for Kansas City Cardiomyopathy Questionnaire (KCCQ). January 19, 2016. Available from: https://www.fda.gov/media/108301/download. Accessed 02 Nov 2021.

Wechalekar AD, Cibeira MT, Gibbs SD, Jaccard A, Kumar S, Merlini G, Palladini G, Sanchorawala V, Schönland S, Venner C, Boccadoro M, Kastritis E. Amyloid. 2023 Mar;30(1):3-17.

Yang H, Wright L, Negishi T, Negishi K, Liu J, Marwick TH. Research to practice: assessment of left ventricular global longitudinal strain for surveillance of cancer chemotherapeutic-related cardiac dysfunction. JACC Cardiovasc Imaging. 2018;11:1196-1201.

Signature Page for VV-CLIN-018963 v1.0

Approval	PPD 15-May-2025 19:20:31 GMT+0000

Approval	PPD 15-May-2025 19:24:49 GMT+0000

Approval	PPD 15-May-2025 21:05:52 GMT+0000

Signature Page for VV-CLIN-018963 v1.0