NOVARTIS

Clinical Development

LNA043

CLNA043A12202 / NCT04864392

A 5-year, randomized, double-blind, placebo-controlled, multi-center study assessing the efficacy, safety, and tolerability of intra-articular regimens of LNA043 versus placebo in patients with symptomatic knee osteoarthritis (ONWARDS)

Statistical Analysis Plan (SAP)

Author:	Trial Statistician
Document type:	SAP Documentation
Document status:	Final
Release date:	04-Sept-2024
Number of pages:	68

Template Version 4.0, Effective from 23-Apr-2021

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)

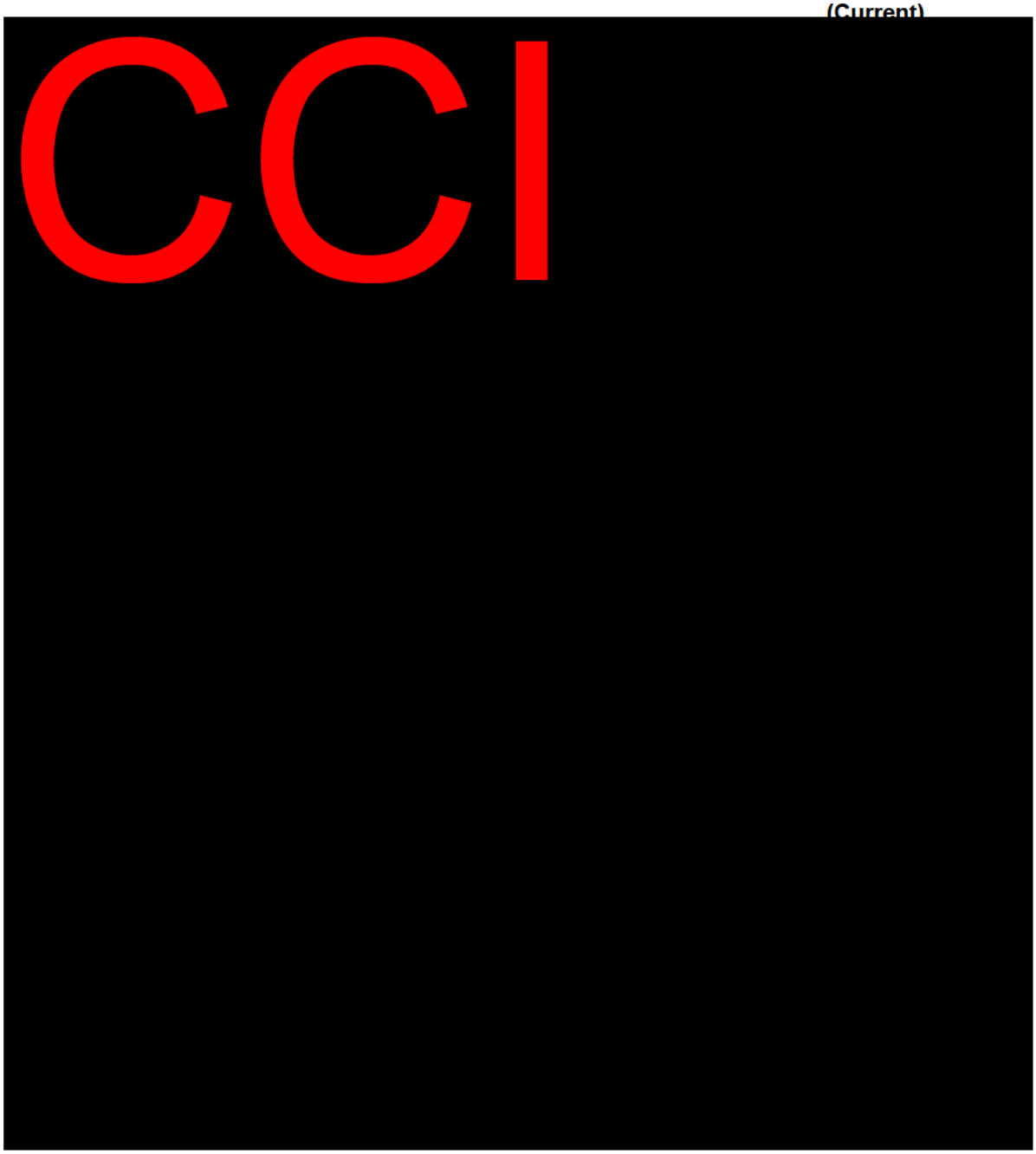


Table of contents

List of tables

List of figures

List of abbreviations

ADA	Anti-Drug Antibodies
AE	Adverse Event
AIR	Acute Inflammatory Reaction
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
ANGPTL3	Angiopoietin-like 3 Protein
AST	Aspartate Aminotransferase
ATC	Anatomical therapeutic chemical
BMI	Body Mass Index
BSL	Baseline
BSP	Biomechanical Sensor Platform
CMQ	Customized MedDRA Query
cMTFC	Central Medial Tibiofemoral Compartment
CRF	Case Report/Record Form (paper or electronic)
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of Variation
DNA	DeoxyriboNucleic Acid
DMC	Data Monitoring Committee
DMS	Document Management System
dSPP	Development Safety Profiling Plan
ECG	Electrocardiogram
eDISH	evaluation of Drug-Induced Serious Hepatotoxicity
EOS	End of Study
EOT	End of Treatment
FAS	Full Analysis Set
FDA	Food and Drug Administration
GGT	Gamma Glutamyl Transferase
HDL	High Density Lipoprotein
HRQL	Health-Related Quality of Life
hsCRP	high sensitivity C-Reactive Protein
i.a.	intra-articular
ICH	International Council for Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use
IG	Immunogenicity
IRT	Interactive Response Technology
ITT	Intent-to-Treat
JSN	Joint Space Narrowing
JSW	Joint Space Width
K-L	Kellgren/Lawrence
KAM	Knee Adduction Moment

KFM	Knee Flexion Moment
kg	Kilogram
LDL	Low Density Lipoprotein
LFT	Liver Function Test
LLN	Lower Limit of Normal
LLOQ	Lower Limit of Quantification
LLQ	Lower Level of Quantification
MAR	Missing At Random
MCID	Minimal Clinically Important Difference
MedDRA	Medical Dictionary for Regulatory Activities
mg	milligram(s)
MI	Multiple Imputation
mL	milliliter(s)
MMRM	Mixed effect Model for Repeated Measurements
CCI	[redacted]
MRI	Magnetic Resonance Imaging
MVPA	Moderate-to-Vigorous Physical Activity
NMQ	Novartis MedDRA Query
NovDTD	Novartis Drug and Therapy Dictionary
NRS	Numerical Rating Scale
NSAIDs	Non-Steroidal Anti-Inflammatory Drug
OA	Osteoarthritis
OARSI	Osteoarthritis Research Society International
OMERACT	Outcome Measures in Rheumatology
PD	Pharmacodynamic(s)
	Protocol Deviation
PGA	Patient Global Assessment
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
PK	pharmacokinetic(s)
PRO	Patient Reported Outcomes
PT	Preferred term
CCI	[redacted]
QoL	Quality of Life
RAS	Randomized Analysis Set
SAE	Serious Adverse Event
SAF	Safety Set
SAP	Statistical Analysis Plan
SAS	Statistical Analysis Software
SD	Standard Deviation
SDC	Smallest Detectable Change
SF-12	Short Form 12 Health Survey

SOC	System Organ Class
SPP	Safety Profiling Plan
SMQ	Standardized MedDRA Query
SS	Symptom Severity
TBL	Total Bilirubin
TFCs	Tibiofemoral Compartments
TKR	Total Knee Replacement
ULN	Upper Limit of Normal
ULQ	Upper Level of Quantification
VKR	Virtual Knee Replacement
vs	Versus
WOMAC	Western Ontario and McMaster Universities Arthritis Index
WPI	Widespread Pain Index

1 Introduction

Data will be analyzed by Novartis according to the data analysis Section 12 of the clinical study protocol. The statistical methodology is described below and any deviations from the protocol are documented. Additional detailed information regarding the analysis methodology is contained in the Appendix section. The analysis plan for DMC will be prepared separately.

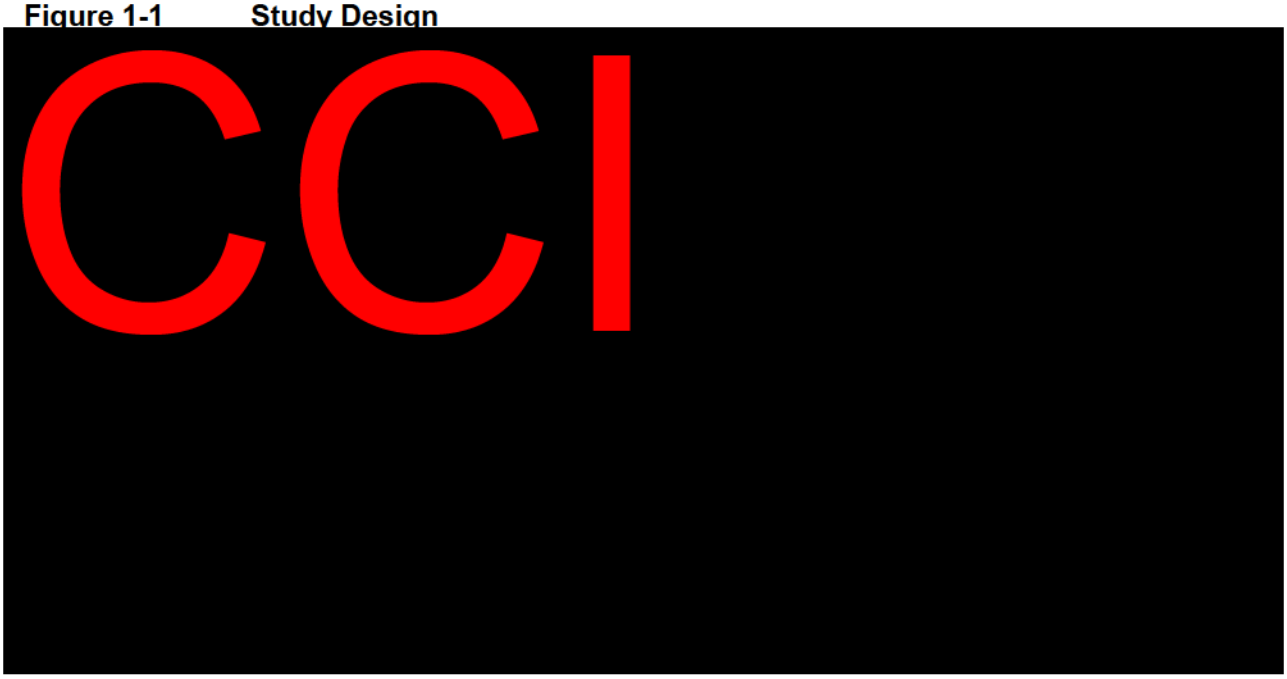
1.1 Study design

CLNA043A12202 is a 2-period, multicenter, randomized, parallel-group, double-blind, placebo-controlled Phase IIb study consisting of a 2-year Core Period, followed by a 3-year Extension Period (with 2-year treatment and 1-year follow-up) aiming to assess the short and long term efficacy, safety and tolerability of multiple intra-articular regimens of LNA043 versus placebo. Optional pre-screening can be considered. Participants with radiographically determined K-L grades 2 or 3 and moderately to severely symptomatic knee OA in the target knee as indicated by a WOMAC pain score of 20-45 on a 0-50 scale prior to initiating study drug were enrolled.

Following the signing of informed consent, participants were assessed for eligibility during a screening period of up to 8 weeks.

On Week 0 (Randomization Visit) of the Core Period, eligible participants were randomized to one of the five treatment groups at a 1:1:1:1:1 allocation. During the entire study, participants will receive CCI mL i.a. injections of either LNA043 CCI mg or CCI mg or placebo (Figure 1-1). Approximately 550 participants were planned to be randomized.

Figure 1-1 Study Design



During the Core Period, all participants will receive i.a. injections of LNA043 or placebo (CCI) every CCI for 2 years (Table 1-1). LNA043 treatment arms will receive active LNA043

CCI mg CCI every CCI , CCI mg CCI every CCI , CCI mg CCI every CCI , or CCI mg CCI every CCI administered i.a. to the target knee. Participants randomized to placebo will receive i.a. injections of saline solution for injection CCI CCI in the Core Period. At the end of the Core Period, participants will enter the Extension Period. During the Extension Period, all participants will receive one i.a. injection of LNA043 or placebo every CCI for 2 years (Table 1-1). Participants who received active LNA043 every CCI in the Core Period will receive one injection of the same dose of active LNA043 correspondingly in the Extension Period (every CCI , respectively). Placebo injections will be used throughout the study in order to keep the same number of injections for all arms and ensure blinding, based on the assigned frequency of treatment. Participants in at least one of these four LNA043 arms may be discontinued from the Extension Period based on efficacy and safety results from the Week 104 primary endpoint database lock. Participants who received placebo during the Core Period will continue to receive placebo every CCI for the Extension Period. The last 1 year of the Extension Period will be a no-treatment follow-up for all arms.

Table 1-1 Summary of study period and treatment

Screening 8 weeks	Core Period 104 weeks	Extension Period 104 weeks
	LNA043 CCI mg i.a. CCI every CCI	LNA043 CCI mg i.a. CCI every CCI
	LNA043 CCI mg i.a. CCI and CCI i.a. CCI every CCI	LNA043 CCI mg i.a. CCI every CCI
	LNA043 CCI mg i.a. CCI followed CCI (after first injection) by CCI i.a. CCI Cycle every CCI	LNA043 CCI mg i.a. CCI followed CCI by CCI i.a. CCI, Cycle every CCI
	LNA043 CCI mg i.a. CCI every CCI	LNA043 CCI mg i.a. CCI every CCI
	Placebo i.a. CCI every CCI	Placebo i.a. CCI every CCI

The primary objective will be to assess the efficacy of each regimen of LNA043 compared to placebo at Week 104 as measured by the mean change from baseline in cartilage thickness of the cMTFC using qMRI of the target knee. An interim analysis will be conducted when CCI
Additional interim analyses may be conducted.

In the event a Public Health emergency as declared by Local or Regional authorities i.e., pandemic, epidemic or natural disaster, that limits or prevents the conduct of on-site study visits, study treatment at home may be considered. Special effort should be made for the EOT/EOS visit. If it is not feasible to conduct the EOT/EOS visit on-site, virtual visits or visits to the participant's home might be considered.

An independent Data Monitoring Committee (DMC) will monitor cumulative semi-blinded safety data during the trial. A separate DMC SAP will specify the analyses to be performed for the DMC reviews.

1.2 Study objectives, endpoints and estimands

Table 1-2 Objectives and related endpoints

Objective(s)	Endpoint(s)
Primary objective(s)	**Endpoint(s) for primary objective(s)**
To evaluate structural changes from baseline in the central medial tibiofemoral compartment (cMTFC) of LNA043 compared with placebo in the target knee at Week 104	Change from baseline in cartilage thickness in the cMTFC at Week 104 assessed by qMRI
Secondary objective(s)	**Endpoint(s) for secondary objective(s)**
To evaluate changes from baseline in OA pain in the target knee of LNA043 compared with placebo at Week 104	Change from baseline at Week 104 in: WOMAC pain WOMAC pain walking on flat surface item
To evaluate changes from baseline in physical function of LNA043 compared with placebo at Week 104	Change from baseline at Week 104 in: WOMAC function
To evaluate structural changes from baseline in the total, medial and lateral tibiofemoral compartments (TFCs) in the target knee of LNA043 compared with placebo at Week 104	Change from baseline in cartilage thickness in the total, medial and lateral TFCs at Week 104 assessed by qMRI
To evaluate changes from baseline in performance-based physical function assessment of LNA043 compared with placebo at Week 104	Change from baseline in physical function at Week 104 40-meter (4×10m) fast-paced walk test 30-second chair stand test 6-minute walking test
To evaluate proportion of patients with structural progression in the target knee of LNA043 compared with placebo using imaging techniques	Proportion of participants demonstrating structural progression at Week 104 defined as: change above the smallest detectable change (SDC) of cartilage thickness by qMRI a loss of medial minimum joint space width (minJSW) ≥ 0.70 mm from baseline by X-ray
To evaluate safety and tolerability of the various LNA043 regimens	Safety and tolerability demonstrated by assessing: Adverse events (AEs) and serious adverse events (SAEs) Incidence of Acute Inflammatory Reactions (AIRs) on the target knee Clinically significant changes in laboratory measures and vital signs, as assessed by the Investigator

	Incidence of binding and neutralizing ADAs in serum
Exploratory objective(s)	**Endpoint(s) for exploratory objective(s)**
To evaluate structural changes in the cMTFC of the target knee over time	Change from baseline in cartilage thickness in the cMTFC assessed by qMRI at Weeks CCI
To evaluate changes from baseline in OA pain in the target knee over time	Change from baseline at Weeks CCI in: WOMAC pain WOMAC pain walking on flat surface item
To evaluate changes from baseline in physical function over time	Change from baseline at Weeks CCI in: WOMAC function
To evaluate structural changes in the total, medial and lateral TFCs in the target knee over time	Change from baseline in cartilage thickness in the total, medial and lateral TFCs assessed by qMRI at Weeks CCI
To evaluate the proportion of patients with structural progression over time in the target knee using imaging techniques	Proportion of participants demonstrating at Week CCI structural progression defined as: change above the SDC of cartilage thickness by qMRI a loss of medial minJSW ≥ 0.70 mm from baseline by X-ray
To evaluate changes in performance-based physical function assessment over time	Change from baseline in physical function at Weeks CCI 40-meter (4×10m) fast-paced walk test 30-second chair stand test 6-minute walking test
To evaluate changes from baseline in total WOMAC score over time	Change from baseline at Weeks CCI in: WOMAC total score
To explore change in disease activity over time	Change from baseline at Weeks CCI in OA disease activity as assessed by Patient Global Assessment (PGA) Change from baseline at Weeks CCI in OA disease activity as assessed by Patient Global Impression of Severity (PGIS) OA disease activity at Weeks CCI as assessed by Patient Global Impression of Change (PGIC)
To assess change in QoL using SF-12v2 over time	Change from baseline at Weeks CCI in QoL using: SF-12v2 physical component summary SF-12v2 mental component summary Total SF-12v2 score
To explore change in CCI as measured by CCI	CCI at Weeks CCI : Change from baseline at Weeks CCI in the following assessments:

	CCI Change from baseline at Weeks CCI in the following assessments: CCI CCI
To evaluate changes in structure over time using X-ray	Change from baseline to Weeks CCI in: fixed-location medial JSW measured on X-ray medial minJSW measured on X-ray
To explore the change from baseline CCI	Change from baseline to CCI
To explore responders using the WOMAC minimal clinically important difference (MCID) over time	Proportion of participants who achieved WOMAC pain MCID CCI Proportion of participants who achieved WOMAC function MCID CCI
To explore responders using the OMERACT-OARSI responder criteria over time	Proportion of OMERACT-OARSI responders at CCI
To evaluate CCI	Change from baseline CCI
To explore the pharmacodynamic effect on cartilage extracellular matrix metabolism over time	Change over time in biomarker levels, including but not limited to: CCI
To evaluate LNA043 pharmacokinetics (PK)	Serum concentrations of LNA043
To evaluate endogenous ANGPTL3 levels	Serum concentrations of ANGPTL3
To explore CCI physical activity and mobility by wrist actigraphy over time	Change from baseline at Weeks CCI in measures of physical activity and mobility.
To explore disease progression by means of TKR and Virtual Knee Replacement (VKR)	Proportion of participants with total knee replacement (TKR) at Week CCI and time to event

	Proportion of participants at Weeks CCI [redacted], fulfilling the following criteria for VKR as defined by: WOMAC Pain ([0-100] scale) + WOMAC Function ([0-100] scale) ≥ 80 points for at least 2 consecutive visits with and without radiographic progression, defined as loss of medial minJSW ≥ 0.50 mm from baseline (Manno et al 2012)
To perform exploratory DNA assessments relating to drug metabolism, cartilage repair, drug target pathway, or other genetic pathways on response (optional)	Exploratory DNA assessments
To explore biomechanical gait parameters by means of a Biomechanical Sensor Platform (BSP) in a subset of participants over time (optional, performed at selected sites only)	Explore changes over time in gait parameters
Explore CCI [redacted]	CCI [redacted]

1.2.1 Primary estimand(s)

The clinical question of interest is: What is the effect of each LNA043 regimen versus placebo on change from baseline in cartilage thickness in the central medial tibiofemoral compartment (cMTFC) of the target knee assessed by qMRI at Week 104 in participants with knee OA who are drug naive or receiving permitted concomitant therapy, had participants taken the assigned treatment for the entire core study duration and regardless of adherence to the allowed period of permitted concomitant therapy?

The justification for targeting this treatment effect is that we wish to estimate the effect of the study drug for the full duration when administered with or without the permitted concomitant therapy.

The primary estimand is described by the following attributes:

1. Population: defined through appropriate inclusion/exclusion criteria to reflect the targeted OA population
2. Primary variable: change from baseline to Week 104 in mean cartilage thickness of the cMTFC of the target knee (CM. Femorotibial Mean Cartilage Thickness tAB (mm))
3. Treatment of interest: the randomized treatment (LNA043 regimens or placebo) taken for the entire core study duration with or without the permitted concomitant therapy

Handling of remaining intercurrent events:

- Treatment discontinuations/disruptions for any reason: had participants taken the assigned treatment for the entire core study duration (hypothetical strategy)
- Use of prohibited medication: ignore (treatment policy strategy)
- Unforeseen non-adherence in the allowed period of permitted concomitant therapies: ignore (treatment policy strategy)

The summary measure: difference in variable means between the treatments.

1.2.2 Secondary estimand(s)

Pain and Function

The clinical questions of interest for secondary objectives pertaining to pain or function are:

- What is the effect of each LNA043 regimen versus placebo on change from baseline in OA pain at Week 104 in participants with knee OA who are drug naive or receiving permitted concomitant therapy, had participants taken the assigned treatment for the entire core study duration and adhered to the allowed period of permitted concomitant therapy?
- What is the effect of each LNA043 regimen versus placebo on change from baseline in function in the target knee at Week 104 in participants with knee OA who are drug naive or receiving permitted concomitant therapy, had participants taken the assigned treatment for the entire core study duration and adhered to the allowed period of permitted concomitant therapy?
- What is the effect of each LNA043 regimen versus placebo on change from baseline in performance-based physical function assessment in the target knee at Week 104 in participants with knee OA who are drug naive or receiving permitted concomitant therapy, had participants taken the assigned treatment for the entire core study duration and adhered to the allowed period of permitted concomitant therapy?

The justification for targeting this treatment effect is that we wish to estimate the effect of the study drug for the full duration when administered with the permitted concomitant therapy.

The estimand definition is described by the following attributes:

1. Population: defined through appropriate inclusion/exclusion criteria to reflect the targeted OA population
2. Variable: change from baseline to Week 104 in the variable of interest
3. Treatment of interest: the randomized treatment (LNA043 regimens or placebo) taken for the entire core study duration with the permitted concomitant therapy

Handling of remaining intercurrent events:

- Treatment discontinuations/disruptions for any reason: had participants taken the assigned treatment for the entire core study duration (hypothetical strategy)
- Use of prohibited medication: had participants taken prohibited medications during the core study duration (hypothetical strategy)

- Unforeseen non-adherence in the allowed period of permitted concomitant therapies: had participants adhered to the allowed period of permitted concomitant therapies (hypothetical strategy)

The summary measure: difference in variable means between the treatments.

Structure

The clinical questions of interest for secondary objectives pertaining to structure are:

- What is the effect of each LNA043 regimen versus placebo on change from baseline in the total, medial and lateral tibiofemoral compartments (TFCs) in the target knee at Week 104 in participants with knee OA who are drug naive or receiving permitted concomitant therapy, had participants taken the assigned treatment for the entire core study duration and regardless of adherence to the allowed period of permitted concomitant therapy?
- What is the effect of each LNA043 regimen versus placebo on the proportion of participants demonstrating structural progression in the target knee at Week 104 in participants with knee OA who are drug naive or receiving permitted concomitant therapy, had participants taken the assigned treatment for the entire core study duration and regardless of adherence to the allowed period of permitted concomitant therapy?

The justification for targeting this treatment effect is that we wish to estimate the effect of the study drug for the full duration when administered with or without adherence to the allowed period of permitted concomitant therapy.

The estimand definition for the secondary objective related to change from baseline is described by the following attributes:

1. Population: defined through appropriate inclusion/exclusion criteria to reflect the targeted OA population
2. Variable: change from baseline to Week 104 in the variable of interest
3. Treatment of interest: the randomized treatment (LNA043 regimens or placebo) taken for the entire core study duration with or without adherence to the allowed period of concomitant therapy

Handling of remaining intercurrent events:

- Treatment discontinuations/disruptions for any reason: had participants taken the assigned treatment for the entire core study duration (hypothetical strategy)
- Use of prohibited medication: ignore (treatment policy strategy)
- Unforeseen non-adherence in the allowed period of permitted concomitant therapies: ignore (treatment policy strategy)

The summary measure: difference in variable means between the treatments.

The estimand definition for the secondary objective related to proportion is described by the following attributes:

1. Population: defined through appropriate inclusion/exclusion criteria to reflect the targeted OA population

2. Variable: proportion of participants demonstrating structural progression at Week 104 in the variable of interest
3. Treatment of interest: the randomized treatment (LNA043 regimens or placebo) taken for the entire core study duration with or without adherence to the allowed period of permitted concomitant therapy

Handling of remaining intercurrent events:

- Treatment discontinuations/disruptions for any reason: had participants taken the assigned treatment for the entire core study duration (hypothetical strategy)
- Use of prohibited medication: ignore (treatment policy strategy)
- Unforeseen non-adherence in the allowed period of permitted concomitant therapies: ignore (treatment policy strategy)

The summary measure: difference in proportions of participants demonstrating structural progression between the treatments.

2 Statistical methods

2.1 Data analysis general information

The analyses will be conducted on all participant data at the time the trial ends and at the time when any interim analyses are performed. Any data analysis carried out independently by the investigator should be submitted to Novartis before publication or presentation.

Summary statistics for continuous variables will include N, mean, standard deviation, minimum, lower quartile, median, upper quartile and maximum. Summary statistics for discrete variables will be presented in the number and percent of participants in each category.

Unless otherwise specified, p-values will be presented as two-sided p-values and two-sided confidence intervals will be displayed. The default level of significance will be set to 5%.

In the core period, efficacy and safety data will be presented by the following treatment groups:

- LNA043 CCI mg CCI , Cycle every CCI
- LNA043 CCI mg CCI followed CCI later (after first injection) by CCI CCI Cycle every CCI
- LNA043 CCI mg CCI , Cycle every CCI
- LNA043 CCI mg ×1 and placebo CCI , Cycle CCI
- Placebo CCI , Cycle every CCI

Participants that are incorrectly dosed will be summarized in their planned treatment arm but will be footnoted in the safety outputs. The cohorts that these participants belong to will also be denoted.

2.1.1 General definitions

2.1.1.1 Study treatment

Study treatment is any drug administered to the study participants as part of the required study procedures, including investigational drug (i.e., LNA043 CCI LNA043 CCI) and control drug (i.e., placebo). See Table 1-1 for details.

No other treatment beyond investigational drug and control drug is included in this study.

2.1.1.2 Study Day 1 and other study days

The first day of administration of randomized study treatment (first dose) is defined as *Study Day 1* or *Day 1*.

All other study days will be labeled relative to Day 1. For event dates on or after Day 1, study day for a particular event date is calculated as [Date of event] – [Date of first dose] + 1, i.e., Day 2, Day 3, etc., will be one day, two days, etc., after Day 1, respectively. For the dates before Day 1, study day for an event date is calculated as [Date of event] – [Date of first dose], i.e., Day -1, Day -2, etc., will be one day, two days, etc., before Day 1, respectively. Duration of an event will be calculated as (Event end date – Event start date + 1).

The descriptor "Day 0" will not be used.

2.1.1.3 Screening, baseline and post-baseline definitions

Screening refers to any procedures (e.g., checking inclusion and exclusion criteria) performed prior to the date of first dose of study treatment or prior to the randomization date. Per protocol, participant informed consent must be obtained prior to performing any study related activity. The date of signing informed consent is the start date of screening period. Any assessment obtained during the screening period will be labeled screening assessment. Assessments made on Day 1 may occur before or after the randomization or the first dose.

For efficacy analyses (except for WOMAC pain, CCI and MRI), baseline is the last assessment (including unscheduled visits) obtained (on or) before the first dose (day). All assessments obtained after the first dose (day) are considered as post-baseline unless otherwise specified.

- For WOMAC pain, baseline is the average of the two WOMAC pain assessments (of the target knee) during the screening period. If only one WOMAC pain assessment is available during the screening period, the single measurement will be used as baseline.
- CCI [redacted]
- For MRI, measurements taken by site (on or) before the first dose (day) or within 4 weeks after the first dose (day) are considered as the baseline. Note that, the target knee is defined as the knee receiving study treatment.
- For actigraphy, a valid day is defined CCI [redacted]

CCI [redacted]

For safety analyses, baseline is the last assessment (including unscheduled visits) obtained (on or) before the first dose (day) of study treatment. All assessments obtained after the first dose (day) of study treatment are considered as post-baseline unless otherwise specified.

If a scheduled baseline assessment value is missing, the screening value will be used instead.

2.1.1.4 Day of last dose of randomized study treatment

The date of last dose will be collected via the CRF.

The participant's exposure will be calculated considering the last dose + 180 or last visit days whichever occurs earlier.

For safety analysis, on-treatment is defined as assessments within last dose + 180 days.

2.1.1.5 End of study / end of treatment

End of study (EOS): the date of the participant's last office visit or other route of follow-up in the study. End of study will be performed at least CCI [redacted] after the last study treatment administration.

End of treatment (EOT): the date when the participant takes their last dose of the study treatment. If the participant discontinues early, then his/her end of treatment visit should be performed at time of discontinuation instead of scheduled trial visit (i.e., Week 104 (EOT1) for Core Treatment Period, Week 208 (EOT2) for Extension Treatment Period). Participants that discontinue early will be excluded from Week 104 (EOT1), 208 (EOT2), and 260 (EOS) visit summaries. Only scheduled visits at Week 104, Week 208 and Week 260 will be summarized,

2.1.1.6 Treatment period

The Core Treatment Period is from the Baseline visit (included) to the end of core treatment visit at Week 104 (included).

The Extension Treatment Period is from the end of core treatment visit (not included) to the end of extension treatment visit at Week 208 (included).

2.1.1.7 Follow-up period

The period from end of treatment (not included) to end of study (included).

For participants who complete the treatment, their follow-up period will be from Week 208 (not included) to Week 260.

For participants who discontinue the treatment, their follow-up period will be from end of treatment (not included) to Week 260 or End of Study if they discontinue study before Week 260.

2.2 Analysis sets

The following analysis sets will be used in this study:

The Randomized Analysis Set (RAS) consists of all randomized participants. Unless otherwise specified, mis-randomized participants (mis-randomized in IRT) will be excluded from the RAS.

Mis-randomized participants are defined as those participants who were mistakenly randomized into the IRT prior to the site confirming all eligibility criteria had been met and to whom no study medication was given. Mis-randomized participants are treated as screen failures.

The Full Analysis Set (FAS) comprises all participants in the RAS to whom study treatment has been assigned by randomization. According to the intent-to-treat principle, participants will be analyzed according to the treatment they have been assigned to during the randomization procedure.

The Safety Set (SAF) includes all participants who received at least one dose of study treatment. Participants will be analyzed according to the study treatment received, where treatment received is defined as the randomized/assigned study treatment if the participant took at least one dose of that treatment or the first treatment received if the randomized/assigned study treatment was never received.

The PK analysis set will include all participants who received study drug, have at least one available valid (i.e., not flagged for exclusion) PK (LNA043) concentration measurement, and for whom there are no protocol deviations with impact on PK data.

The ANGPTL3 analysis set will include all participants who have at least one available valid (i.e., not flagged for exclusion) ANGPTL3 concentration measurement.

The Immunogenicity Prevalence Set includes all participants in the Safety Set with a non-missing baseline ADA sample or at least one non-missing post-baseline ADA sample.

The Immunogenicity incidence set includes all participants in the Immunogenicity prevalence set with a non-missing baseline ADA sample and at least one non-missing post-baseline ADA sample.

2.2.1 Subgroup of interest

Additional subgroup analysis may be conducted if deemed necessary.

Subgroup analyses for the Japanese subset will be defined in a separate analysis plan.

2.3 Patient disposition, demographics and other baseline characteristics

2.3.1 Patient disposition

The number of participants screened will be presented. In addition, the reasons for screen failures will be provided. The number and percentage of participants in the randomized set who completed the study periods and who discontinued the study prematurely (including the reason for discontinuation) will be presented at the end of each treatment period (Week 104 and Week 208) and follow-up period, if appropriate, for each treatment group and all participants.

"Off-treatment" is defined as those participants that have discontinued study treatment but have remained in the study and have had subsequent visits/assessments with in within the study period, not including the End of Study visit.

For each protocol deviation (PD), the number and percentage of participants for whom the PD applies will be tabulated. Protocol deviations due to COVID-19 will be reported in separate categories.

Additional analysis due to COVID-19 might be provided, e.g., summary of participants missed visit and/or treatment due to COVID-19.

2.3.2 Demographics and other baseline characteristics

The following common background and demographic variables will be summarized:

Continuous variables:

- Age
- Height
- Weight
- Body mass index (BMI) = (body weight in kilograms) / (height in meters)^2

For BMI, height and body weight, the last value prior to randomization is used. If there is no weight recorded prior to administration of study drug, BMI will be missing.

Categorical variables:

- Sex
- Race
- Ethnicity

Baseline disease characteristics will also be summarized for the following variables:

- WOMAC pain, WOMAC function, CCI [redacted], patient global assessment of pain, widespread pain index, symptom severity scale, patient global impression of severity , physical function (40-meter fast-paced walk test, 30-second chair stand test, 6-minute walking test), hsCRP (mg/L), time since first diagnosis of knee OA (years), joint space narrowing (JSN) by X-ray, joint space width (JSW) by X-ray, K-L grade, cartilage thickness by qMRI.

Time since diagnosis of knee OA will be calculated using the following formula:

Time since diagnosis = (inform consent date –first diagnosis date + 1)/365.25

The first diagnosis date of knee OA will be imputed according to the imputation rules in Appendix 5.1.4.

Unless otherwise specified, summary statistics will be presented for continuous variables for each treatment group and for all participants (total) in the randomized set. The number and percentage of participants in each category will be presented for categorical variables for each treatment group and all participants (total) in the randomized set.

2.3.3 Medical history

Any condition entered on the *Relevant medical history* CRF will be coded using the MedDRA dictionary. They will be summarized by system organ class (SOC) and preferred term (PT) of the MedDRA dictionary. Summaries for knee OA specific medical history will be provided. Listings of medical history including the reported term/verbatim will be provided as well.

Unless otherwise specified, analyses will be based on the randomized set.

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

2.4.1 Study treatment / compliance

The analysis of study treatment data will be based on the safety set.

The number of visits with active and placebo injections received will be presented by treatment group. The duration of exposure to study treatment will also be summarized by treatment group. In addition, the total dose administered will be presented by treatment group. Total dose administrated is defined as the sum of the injection dosage across all visits. Total dosage will be defined in mg.

Duration of exposure will be defined as the time from first dose of study treatment to the end of study period. The end of study period will be defined as the last dose + 180 days or last visit whichever occurs earlier. For participants who discontinued and have their last visit earlier than last dose + 180 days, the end of study exposure will be the date of the last study visit in the corresponding treatment period.

Duration of exposure (days) = min (‘end of study period’ date, last dose date + 180 days) - first dose date + 1

Duration of exposure (years) = duration of exposure (days) / 365.25

Duration of exposure (100 subject years) = duration of exposure (years) / 100

The analyses of duration of exposure described above will be done for the entire study treatment period.

2.4.1.1 Analysis visit window

No analysis visit windows will be defined for this study. Consequently, all by-visit summaries will be performed as per below scheduled visits.

Table 2-1 CCI

2.4.2 Prior, concomitant and post therapies

Medications will be identified using the Novartis Drug and Therapy Dictionary (NovDTD) including Anatomical Therapeutic Chemical (ATC) code. Prior and concomitant treatments will be summarized in separate tables by treatment group for the safety set. Concomitant treatments will be displayed by study period as appropriate. For example, medications started in the follow-up periods might be summarized separately.

Prior medications are defined as treatments taken and stopped prior to first dose of study treatment. Any medication given at least once between the day of first dose of randomized study treatment and within 180 days after last dose will be a concomitant medication, including those which were started pre-baseline and continued into the period where study treatment is administered.

Prior and concomitant medications will be summarized by treatment group in separate tables. Medications will be presented in alphabetical order, by Anatomical Therapeutic Classification (ATC) codes and grouped by anatomical main group (the 1st level of the ATC codes). Tables will show the overall number and percentage of participants receiving at least one treatment of a particular ATC code and at least one treatment in a particular anatomical main group.

Prior non-drug therapies and procedures are defined as non-drug therapies and procedures done prior to first dose of study treatment. Any non-drug therapies and procedures done between the day of first dose of study treatment and within 180 days after last dose will be defined as concomitant non-drug therapies and procedures, including those which were started pre-baseline and continued into the period where study treatment is administered.

Significant prior and concomitant non-drug therapies and procedures will be summarized by primary system organ class and MedDRA preferred term.

The number and percentage of participants receiving prior and concomitant knee OA therapy will be presented by randomized treatment group and the total duration of previous exposure to knee OA therapies.

The number and percentage of participants receiving concomitant NSAIDs and pain medications will be summarized.

Prior or concomitant medication will be identified by comparing recorded or imputed start and end dates of medication taken to the reference start date.

2.5 Analysis supporting primary objective(s)

2.5.1 Primary endpoint(s)

The primary efficacy variable is change from baseline in mean cartilage thickness in the cMTFC at Week 104 determined by qMRI of the target knee in participants with knee OA. The analysis of the primary variable will be based on the FAS population.

2.5.2 Statistical hypothesis, model, and method of analysis

The null statistical hypothesis being tested is that LNA043 regimens are not superior to Placebo in the mean change from baseline in mean cartilage thickness of the cMTFC at Week 104.

Let μ_j denote the mean change from baseline in mean cartilage thickness of the cMTFC at Week 104 for treatment regimens j, $j = 0, 1, 2, 3$ or 4 where

- 0 corresponds to Placebo
- 1 corresponds to LNA043 CCI mg CCI Cycle every CCI
- 2 corresponds to LNA043 CCI mg CCI and CCI Cycle every CCI
- 3 corresponds to LNA043 CCI mg CCI followed CCI later (after first injection) by CCI, Cycle every CCI
- 4 corresponds to LNA043 CCI mg CCI, Cycle every CCI

In statistical terms, H_j: $\mu_j \leq \mu_0$, H_{Aj}: $\mu_j > \mu_0$,

H_1: LNA043 CCI mg CCI, Cycle every CCI is not superior to Placebo with respect to mean change from baseline in cartilage thickness of the cMTFC at Week 104

H_2: LNA043 CCI mg CCI and CCI Cycle every CCI is not superior to Placebo with respect to mean change from baseline in cartilage thickness of the cMTFC at Week 104

H_3: LNA043 CCI mg CCI followed CCI later (after first injection) by CCI Cycle every CCI is not superior to Placebo with respect to mean change from baseline in cartilage thickness of the cMTFC at Week 104

H_4: LNA043 CCI mg CCI Cycle every CCI is not superior to Placebo with respect to mean change from baseline in cartilage thickness of the cMTFC at Week 104

The primary analysis will be conducted using mixed-effect model repeated measures (MMRM) which is valid under the missing at random (MAR) assumption, with treatment and analysis

visit as factors and baseline score as a covariate in the model. Treatment by analysis visit and baseline score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for this model.

The significance of the treatment effects for LNA043 regimens at Week 104 will be determined from the pairwise comparisons using the Dunnett test performed between LNA043 regimens and placebo at $\alpha = 0.05$ (one-sided, family-wise type-I-error).

A graph of the estimated mean change from baseline in mean cartilage thickness of the cMTFC by visit and treatment group will also be presented.

2.5.3 Handling of intercurrent events

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

- Treatment discontinuations/disruptions for any reason: data will be censored after treatment discontinuation or first disruption for the primary analysis (hypothetical strategy)
- Use of prohibited medication: Data collected after use of prohibited medications will be used for the primary analysis (treatment policy strategy)
- Unforeseen non-adherence in the allowed period of permitted concomitant therapies: data collected after non-adherence in the allowed period of permitted concomitant therapies will be used for the primary analysis (treatment policy strategy)

If a patient has multiple intercurrent events, we will only consider the first intercurrent event that occurred and that used the hypothetical strategy. The data censored by hypothetical strategy will be imputed by MMRM under a MAR assumption.

Missing data related to the treatment policy will be implicitly imputed by MMRM under a MAR assumption.

For analyses, if all post-baseline values are missing then these missing values will not be imputed and this participant will be removed from the analysis, i.e., it might be that the number of participants providing data to the analysis is smaller than the number of participants in the FAS.

2.5.4 Handling of missing values not related to intercurrent event

The missing data not related to the intercurrent events will be implicitly imputed by MMRM under a MAR assumption.

2.5.5 Sensitivity analyses

The following sensitivity analyses will be performed for the primary estimand, to assess the robustness of the estimation in the presence of deviations from the assumptions specified in the primary analysis.

Tipping point analysis:

The same mixed model for repeated measures (MMRM) model as for the primary estimation will be adopted. Intercurrent events will be handled in the same way as in the primary analysis.

Subsequently, for the active treatment arms these imputed values will be further worsened via the application of increasingly large penalties specified by a sensitivity parameter (i.e., delta) according to the tipping point method (Permutt 2016). The same model as in the primary analysis will be fitted and p-values of the Dunnett test will be saved. With all combinations of delta and p-values, the one such that the Dunnett test is insignificant will be summarized.

Impact of MRI scanner changes between visits on the primary endpoint:

An additional sensitivity analysis will be conducted by excluding visits from participants who had a documented scanner change. The same model from the primary analysis will be adopted .

2.5.6 Supplementary analyses

The target population, the primary variable and the summary measure for the supplementary estimand are the same as for the primary estimand. Differently from the primary estimand, the intervention effect for this supplementary estimand is the effect in a treatment policy scenario where participants would not be able to always adhere to study treatment regimen or the use of prohibited medication.

The estimation method is the same as for the primary estimand except that data after study treatment discontinuation/disruption or use of prohibited medication for any reason in all treatment arms will be included and real missing data will be imputed in the MMRM.

2.6 Analysis supporting secondary objectives

2.6.1 Secondary endpoint(s)

The secondary efficacy variables include,

- Change from baseline in WOMAC pain (on a normalized 0-100 scale) at Week 104
- Change from baseline in WOMAC pain walking on flat surface (on a normalized 0-100 scale) at Week 104
- Change from baseline in WOMAC function (on a normalized 0-100 scale) at Week 104
- Change from baseline in cartilage thickness in the total, medial and lateral TFCs in the target knee at Week 104
- Change from baseline in physical function (40-meter fast-paced walk test, 30-second chair stand test, 6-minute walking test) at Week 104
- Proportion of participants demonstrating structural progression at Week 104

The analysis for all secondary efficacy endpoints will be based on the FAS population.

2.6.2 Statistical hypothesis, model, and method of analysis

The secondary efficacy variables and the hypotheses are described below. Secondary efficacy variables will be analyzed using the FAS population.

The following null hypotheses will be included in the testing strategy, and type-I-error will be set such that a family-wise type-I-error of 5% (one-sided) will be controlled for at each individual endpoint using a Dunnett's adjustment.

WOMAC pain at Week 104

H_5: LNA043 CCI mg CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC pain at Week 104

H_6: LNA043 CCI mg × 1 and placebo CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC pain at Week 104

H_7: LNA043 CCI mg CCI followed CCI later (after first injection) by Placebo CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC pain at Week 104

H_8: LNA043 CCI mg CCI, Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC pain at Week 104

WOMAC pain will be rescaled to a normalized 0-100 scale.

Between-treatment differences in the change from baseline in WOMAC pain will be evaluated using MMRM which is valid under the MAR assumption, with treatment group and analysis visit as factors and baseline WOMAC pain as a continuous covariate. Treatment by analysis visit and baseline WOMAC pain by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model.

The significance of the treatment effects for LNA043 regimens at Week 104 will be determined from the pairwise comparisons using the Dunnett test performed between LNA043 regimens and placebo at $\alpha = 0.05$ (one-sided, family-wise type-I-error).

WOMAC pain walking on flat surface at Week 104

H_9: LNA043 CCI mg CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC pain walking on flat surface at Week 104

H_{10}: LNA043 CCI mg × 1 and placebo CCI, Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC pain walking on flat surface at Week 104

H_{11}: LNA043 CCI mg CCI followed CCI later (after first injection) by Placebo CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC pain walking on flat surface at Week 104

H_{12}: LNA043 CCI mg CCI, Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC pain walking on flat surface at Week 104

WOMAC pain walking on flat surface will be rescaled to a normalized 0-100 scale.

Between-treatment differences in the change from baseline in WOMAC pain walking on flat surface will be evaluated using MMRM which is valid under the MAR assumption, with treatment group and analysis visit as factors and baseline WOMAC pain walking on flat surface

as a continuous covariate. Treatment by analysis visit and baseline WOMAC pain walking on flat surface by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model.

The significance of the treatment effects for LNA043 regimens at Week 104 will be determined from the pairwise comparisons using the Dunnett test performed between LNA043 regimens and placebo at $\alpha = 0.05$ (one-sided, family-wise type-I-error).

WOMAC function at Week 104

H_{13}: LNA043 CCI mg CCI , Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC function at Week 104

H_{14}: LNA043 CCI mg × 1 and placebo CCI , Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC function at Week 104

H_{15}: LNA043 CCI mg CCI followed CCI later (after first injection) by Placebo CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC function at Week 104

H_{16}: LNA043 CCI mg CCI , Cycle CCI is not superior to Placebo with respect to mean change from baseline in WOMAC function at Week 104

WOMAC function will be rescaled to a normalized 0-100 scale.

Between-treatment differences in the change from baseline in WOMAC function will be evaluated using MMRM which is valid under the MAR assumption, with treatment group and analysis visit as factors and baseline WOMAC function score as a continuous covariate. Treatment by analysis visit and baseline WOMAC function score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model.

The significance of the treatment effects for LNA043 regimens at Week 104 will be determined from the pairwise comparisons using the Dunnett test performed between LNA043 regimens and placebo at $\alpha = 0.05$ (one-sided, family-wise type-I-error).

Cartilage thickness in the total TFCs at Week 104

H_{17}: LNA043 CCI mg CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in total TFCs assessed by qMRI at Week 104

H_{18}: LNA043 CCI mg × 1 and placebo CCI , Cycle CCI is not superior to Placebo with respect to mean change from baseline in total TFCs assessed by qMRI at Week 104

H_{19}: LNA043 CCI mg CCI followed CCI later (after first injection) by Placebo CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in total TFCs assessed by qMRI at Week 104

H_{20}: LNA043 CCI mg CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in total TFCs assessed by qMRI at Week 104

Between-treatment differences in the change from baseline in cartilage thickness in the total, TFCs in the target knee will be evaluated using MMRM which is valid under the MAR assumption, with treatment group and analysis visit as factors and baseline score as a continuous covariate. Treatment by analysis visit and baseline score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model.

The significance of the treatment effects for LNA043 regimens at Week 104 will be determined from the pairwise comparisons using the Dunnett test performed between LNA043 regimens and placebo at $\alpha = 0.05$ (one-sided, family-wise type-I-error).

Cartilage thickness in the medial TFCs at Week 104

H_{21}: LNA043 CCI mg CCI , Cycle CCI is not superior to Placebo with respect to mean change from baseline in medial TFCs assessed by qMRI at Week 104

H_{22}: LNA043 CCI mg × 1 and placebo CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in medial TFCs assessed by qMRI at Week 104

H_{23}: LNA043 CCI mg CCI followed CCI later (after first injection) by Placebo CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in medial TFCs assessed by qMRI at Week 104

H_{24}: LNA043 CCI mg CCI , Cycle CCI is not superior to Placebo with respect to mean change from baseline in medial TFCs assessed by qMRI at Week 104

Between-treatment differences in the change from baseline in cartilage thickness in the medial TFCs in the target knee will be evaluated using MMRM which is valid under the MAR assumption, with treatment group and analysis visit as factors and baseline score as a continuous covariate. Treatment by analysis visit and baseline score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model.

The significance of the treatment effects for LNA043 regimens at Week 104 will be determined from the pairwise comparisons using the Dunnett test performed between LNA043 regimens and placebo at $\alpha = 0.05$ (one-sided, family-wise type-I-error).

Cartilage thickness in the lateral TFCs at Week 104

H_{25}: LNA043 CCI mg CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in lateral TFCs assessed by qMRI at Week 104

H_{26}: LNA043 CCI mg × 1 and placebo CCI , Cycle CCI is not superior to Placebo with respect to mean change from baseline in lateral TFCs assessed by qMRI at Week 104

H_{27}: LNA043 CCI mg CCI followed CCI later (after first injection) by Placebo CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in lateral TFCs assessed by qMRI at Week 104

H_{28}: LNA043 CCI mg CCI , Cycle CCI is not superior to Placebo with respect to mean change from baseline in lateral TFCs assessed by qMRI at Week 104

Between-treatment differences in the change from baseline in cartilage thickness in the lateral TFCs in the target knee will be evaluated using MMRM which is valid under the MAR assumption, with treatment group and analysis visit as factors and baseline score as a continuous covariate. Treatment by analysis visit and baseline score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model.

The significance of the treatment effects for LNA043 regimens at Week 104 will be determined from the pairwise comparisons using the Dunnett test performed between LNA043 regimens and placebo at $\alpha = 0.05$ (one-sided, family-wise type-I-error).

40-meter (4×10m) fast-paced walk test at Week 104

H_{29}: LNA043 CCI mg CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 40-meter (4×10m) fast-paced walk test at Week 104

H_{30}: LNA043 CCI mg × 1 and placebo CCI, Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 40-meter (4×10m) fast-paced walk test at Week 104

H_{31}: LNA043 CCI mg CCI followed CCI later (after first injection) by Placebo CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 40-meter (4×10m) fast-paced walk test at Week 104

H_{32}: LNA043 CCI mg CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 40-meter (4×10m) fast-paced walk test at Week 104

Between-treatment differences in the change from baseline in 40-meter fast-paced walk test will be evaluated using MMRM which is valid under the MAR assumption, with treatment group and analysis visit as factors and baseline score as a continuous covariate. Treatment by analysis visit and baseline score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model.

The significance of the treatment effects for LNA043 regimens at Week 104 will be determined from the pairwise comparisons using the Dunnett test performed between LNA043 regimens and placebo at $\alpha = 0.05$ (one-sided, family-wise type-I-error).

30-second chair stand test at Week 104

H_{33}: LNA043 CCI mg CCI, Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 30-second chair stand test at Week 104

H_{34}: LNA043 CCI mg × 1 and placebo CCI, Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 30-second chair stand test at Week 104

H_{35}: LNA043 CCI mg CCI followed CCI later (after first injection) by Placebo CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 30-second chair stand test at Week 104

H_{36}: LNA043 CCI mg CCI, Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 30-second chair stand test at Week 104

Between-treatment differences in the change from baseline in 30-second chair stand test will be evaluated using MMRM which is valid under the MAR assumption, with treatment group and analysis visit as factors and baseline score as a continuous covariate. Treatment by analysis visit and baseline score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model.

The significance of the treatment effects for LNA043 regimens at Week 104 will be determined from the pairwise comparisons using the Dunnett test performed between LNA043 regimens and placebo at $\alpha = 0.05$ (one-sided, family-wise type-I-error).

6-minute walking test at Week 104

H_{37}: LNA043 CCI mg CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 6-minute walking test at Week 104

H_{38}: LNA043 CCI mg × 1 and placebo CCI, Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 6-minute walking test at Week 104

H_{39}: LNA043 CCI mg CCI followed CCI later (after first injection) by Placebo CCI Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 6-minute walking test at Week 104

H_{40}: LNA043 CCI mg CCI, Cycle CCI is not superior to Placebo with respect to mean change from baseline in physical function assessed by 6-minute walking test at Week 104

Between-treatment differences in the change from baseline in 6-minute walking test will be evaluated using MMRM which is valid under the MAR assumption, with treatment group and analysis visit as factors and baseline score as a continuous covariate. Treatment by analysis visit and baseline score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model.

The significance of the treatment effects for LNA043 regimens at Week 104 will be determined from the pairwise comparisons using the Dunnett test performed between LNA043 regimens and placebo at $\alpha = 0.05$ (one-sided, family-wise type-I-error).

Loss of medial minimum joint space width (minJSW) ≥ 0.70 mm (structural progression) from baseline by X-ray at Week 104

H_{45}: LNA043 CCI mg CCI, Cycle CCI is not superior to Placebo with respect to proportion of participants demonstrating structural progression in the target knee defined as a

loss of medial minimum joint space width (minJSW) ≥ 0.70 mm from baseline by X-ray at Week 104

H_{46}: LNA043 CCI mg × 1 and placebo CCI , Cycle CCI is not superior to Placebo with respect to proportion of participants demonstrating structural progression in the target knee defined as a loss of medial minimum joint space width (minJSW) ≥ 0.70 mm from baseline by X-ray at Week 104

H_{47}: LNA043 CCI mg CCI followed CCI later (after first injection) by Placebo CCI Cycle CCI is not superior to Placebo with respect to proportion of participants demonstrating structural progression in the target knee defined as a loss of medial minimum joint space width (minJSW) ≥ 0.70 mm from baseline by X-ray at Week 104

H_{48}: LNA043 CCI mg CCI Cycle CCI is not superior to Placebo with respect to proportion of participants demonstrating structural progression in the target knee defined as a loss of medial minimum joint space width (minJSW) ≥ 0.70 mm from baseline by X-ray at Week 104

Multiple imputation (MI) approach under MAR assumptions will be applied to handle missing data for structural progression.

The proportion of participants demonstrating structural progression will be evaluated using a logistic regression model with treatment group as a factor. Difference in marginal response proportions with p-value and 95% confidence interval (CI) will be presented.

The significance of the treatment effects for LNA043 regimens at Week 104 will be determined from the pairwise comparisons using the Dunnett test performed between LNA043 regimens and placebo at $\alpha = 0.05$ (one-sided, family-wise type-I-error).

2.6.3 Handling of intercurrent events

Pain and Function

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

- Treatment discontinuations/disruptions for any reason: data will be censored after treatment discontinuation or first disruption (hypothetical strategy)
- Use of prohibited medications: data will be censored after first use of prohibited medication (hypothetical strategy)
- Unforeseen non-adherence in the allowed period of permitted concomitant therapies: data will be set to missing on the visit of non-adherence in the allowed period of permitted concomitant therapies (hypothetical strategy)

The MMRM model implicitly imputes missing data under a missing at random assumption. For analyses, if all post-baseline values are missing then these missing values will not be imputed and this participant will be removed from the analysis, i.e., it might be that the number of participants providing data to the analysis is smaller than the number of participants in the FAS.

Imputation under MAR relies on the assumption that unbiased estimates can be obtained by borrowing information from participants with collected data that are similar in regards to model baseline covariates and measurements collected at prior visits.

Structure

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

- Treatment discontinuations/disruptions for any reason: data will be censored after treatment discontinuation or first disruption (hypothetical strategy)
- Use of prohibited medication: Data collected after use of prohibited medications will be used for analysis (treatment policy strategy)
- Unforeseen non-adherence in the allowed period of permitted concomitant therapies: data collected after non-adherence in the allowed period of permitted concomitant therapies will be used (treatment policy strategy)

The data censored by hypothetical strategy will be handled as follows:

- For continuous efficacy variables (e.g., cartilage thickness in the total, etc.), the MMRM model implicitly imputes missing data under the MAR assumption.
- For binary efficacy variables (e.g., structural progression), the MI approach imputes missing data under the MAR assumption.

2.6.4 Handling of missing values not related to intercurrent event

The missing data not related to the intercurrent events will be handled as follows:

- For continuous efficacy variables (e.g., cartilage thickness in the total, etc.), the MMRM model implicitly imputes missing data under the MAR assumption.
- For binary efficacy variables (e.g., structural progression), the MI approach imputes missing data under the MAR assumption.

2.6.5 Sensitivity analyses

Not applicable.

2.6.6 Supplementary analyses

Not applicable.

2.7 Safety analyses

All safety analyses will be based on the SAF and performed on treatment received or actual treatment as described below:

The actual treatment or treatment received for summaries of safety data will differ to the treatment assigned at randomization only if a participant received the wrong treatment during the entire study.

For those participants who did not receive the randomized treatment, i.e., who received erroneously the wrong treatment at least once, an additional AE listing will be prepared displaying which events occurred after the treatment errors.

In addition, for participants who discontinue study treatment but continue with study participation, an additional AE listing will be prepared displaying which events occurred after the study treatment discontinuation.

Summaries may be performed separately for core treatment period and entire treatment period (including follow-up).

Safety summaries (tables, figures) will include only data from the on-treatment period with the exception of baseline data which will also be summarized where appropriate (e.g., change from baseline summaries). In addition, a separate summary for death including on treatment and post treatment deaths will be provided. In particular, summary tables for adverse events (AEs) will summarize only on-treatment events, with a start date during the on-treatment period (treatment-emergent AEs).

The on-treatment period lasts from the date of first administration of study treatment up to last administration + 180 days.

2.7.1 Adverse events (AEs)

All reported AEs will be coded based on the latest MedDRA version practically possible (version 27.0 or above) and clarified in footnotes of reports.

The number (and percentage) of participants with treatment emergent adverse events will be summarized in the following ways:

- by treatment, primary system organ class
- by treatment, preferred term
- by treatment, primary system organ class and preferred term.
- by treatment, primary system organ class, preferred term and maximum severity.

Confidence intervals for the crude rate will be derived as described in Section 5.5.3.

Treatment emergent adverse events are defined as events started on or after the first dose of study medication or events present prior to the first dose of study medication but increased in severity on or after dosing based on preferred term and within last dose + 180 days (inclusive).

A participant with multiple adverse events within a primary system organ class is only counted once towards the total of the primary system organ class.

Separate summaries will be provided for study treatment related adverse events, death, serious adverse event, other significant adverse events leading to discontinuation and adverse events leading to dose adjustment or interruption.

The number (and proportion) of participants with adverse events of special interest will be summarized by treatment.

Adverse events reported will be presented in descending frequency according to its incidence in the total LNA043 group (combining all LNA043 treatment groups) starting from the most

common event. AEs will be summarized by presenting - for each treatment group (including any LNA043) - the number and percentage of participants having any AE, having an AE in each primary system organ class and having each individual AE (preferred term).

Summaries will also be presented for AEs by severity and for study treatment related AEs. If a particular AE 'severity' is missing, this variable will be listed as missing and treated as missing in summaries. If a participant reported more than one adverse event with the same preferred term, the adverse event with the greatest severity will be presented. If a participant reported more than one adverse event within the same primary system organ class, the participant will be counted only once with the greatest severity at the system organ class level, where applicable. Of note, AEs by severity will be provided by SOC in the study report and AEs by severity and PTs may be provided if required at ad-hoc basis.

Adverse events will also be reported separately by customized MedDRA queries (CMQ/NMQ). The MedDRA version (version 27.0 or above) used for reporting the study will be described in a footnote.

The most common adverse events reported ($\geq z$% in any group for each preferred term in the SOC-PT table or $\geq z$% in any group for each grouping term table) will be presented in descending frequency according to its incidence in total LNA043 group starting from the most common event. Here threshold value z is set to 2 (%) but it may be updated following review of the dry run outputs.

In addition, exposure time-adjusted incidence rates including 95% confidence intervals will be provided for the entire treatment period following the guideline as below:

- Primary SOC level for AE and SAE
- Level 1 for risks
- PT level for SAE
- PT level for the most common AEs ($\geq$ 2% incidence) in the any LNA043 group
- Other selected AEs on lower levels (e.g., PT), if appropriate

A graphical display of the crude rates or exposure adjusted-rates within system organ classes and relative risks, if appropriate, will be presented as follows:

For all AEs regardless of severity and seriousness, the point estimate (i.e., relative frequency for evaluation of the core treatment period and exposure adjusted incidence for evaluation of entire treatment/study) within system organ classes will be presented graphically with system organ class on the y-axis. This figure will consist of two panels:

1. point estimate of AEs,
2. point estimate of serious AEs.

The placebo group will be displayed with a bar whereas dots will be used for LNA043 treatment groups.

If appropriate, additional plots will be provided showing point estimates and confidence intervals on the left panel and numeric values of point estimate and confidence interval on the right panel of the figure. This will be done separately for all AEs and all SAEs.

Algorithms for date imputations will be provided in Section 5.1.2.

A listing of non-treatment emergent adverse events will be provided. These adverse events occurred before the first dose of the study treatment or after the last dose + 180 days. The crude incidence rate will be provided without treatment information.

For SAEs occurred during screening a listing will be prepared for all participants screened including screening failures.

An overview of the safety analyses (i.e., crude incidence, exposure time-adjusted incidence) which will be performed for adverse events and other binary safety variables (e.g., labs and vital signs) for each analysis period is described in Table 2-2.

Table 2-2 Overview of analyses on some safety endpoints

Analysis period	AEs & SAEs	SPP risks	AEs by severity	Study drug related AEs	Notables for (vitals/ ECG), lab criteria
Day 1- Week 104	• crude incidence	• crude incidence	• crude incidence	• crude incidence	• crude incidence
Entire treatment	• crude incidence • exposure time adjusted incidence	• crude incidence • exposure time adjusted incidence	• crude incidence	• crude incidence	• crude incidence

2.7.1.1 Disclosure reporting

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables on treatment emergent AEs: non-SAEs with an incidence greater than 5%, and deaths and SAEs including events suspected to be related to study treatment, will be provided by treatment, SOC and PT on the Safety Set.

If, for the same patient, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- a single occurrence will be counted if there is ≤ 1 day gap between the end date of the preceding AE and the start date of the consecutive AE
- more than one occurrence will be counted if there is > 1 day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE / SAE suspected to be related to study treatment/ non-SAE has to be checked in a block e.g. among AE's in a ≤ 1 day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

The number of deaths resulting from SAEs suspected to be related to study treatment and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

2.7.1.2 Adverse events of special interest / grouping of AEs

Crude incidence rate and exposure adjusted incidence rates for adverse events of special interest may be summarized by treatment if required.

Crude rate of risks from Case Retrieval Sheet will be provided for all (non-serious and serious) cases and for all serious cases. Exposure-time adjusted rates will be provided for treatment period including all data for all (non-serious and serious) cases and for all serious cases. In addition, listings will be provided for the related AE risks.

Risk measures and confidence intervals will be derived according to Section 5.5.

The version of the Case Retrieval Sheet used for the analyses will be described in a footnote. This includes MedDRA version and Novartis MedDRA Query (NMQ) dictionary date.

For further background information and medical questions to be addressed please refer to the Development Safety Profiling Plan stored in the Document Management System (DMS).

Of note, for the evaluation of risks primary and secondary system organ classes of the MedDRA dictionary will be considered.

2.7.2 Deaths

Separate summary and listing will be provided for deaths.

2.7.3 Laboratory data

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

The general guideline for laboratory summaries (including ECGs and Vital signs in Section 2.7.4.1 and Section 2.7.4.2) are as below:

- All the summary of lab outputs (newly occurring notables, maximum changes, shift tables, by visit summary statistics) will consider the "on-treatment" data. i.e., all assessments up to last dose + 180 days.
- Follow up visit summary: using CRF visits (including early discontinued participants). i.e., not apply visit window and cut-off day 180. Summary of follow up visit outputs may be provided if required. The listing will provide follow up visit records.
- All records are displayed in the listing with the on-treatment flag. i.e., occurred up to last dose + 180 days- yes or no, as well as eCRF visits and epoch.

The summary of laboratory evaluations will be presented for three groups of laboratory tests (hematology, clinical chemistry and urinalysis). In addition to the individual laboratory parameters the ratios “total cholesterol / HDL” will be derived and summarized.

For urinalysis, frequency tables will be presented except for Specific Gravity and pH. Change from baseline will be presented for these parameters.

Descriptive summary statistics for the change from baseline to each study visit will be presented. These descriptive summaries will be presented by laboratory test and treatment group. Change from baseline will only be summarized for participants with both baseline and post baseline values and will be calculated as:

change from baseline = post baseline value – baseline value

Reported laboratory assessments with either a less than or greater than sign ("<" or ">") will be used for analysis after removal of the sign and conversion to standard unit. These laboratory values will be displayed in listings using the standard unit with the reported sign ("<" or ">"). For each parameter, the maximum change (maximum decrease and maximum increase) from baseline within treatment period will be analyzed analogously.

In addition, shift tables may be provided as required at ad-hoc basis for all parameters to compare a participant's baseline laboratory evaluation relative to the visit's observed value. For the shift tables, the normal laboratory ranges will be used to evaluate whether a particular laboratory test value was normal, low, or high for each visit value relative to whether or not the baseline value was normal, low, or high. The shifts to the most extreme laboratory test value within a treatment phase (either core or entire) will be presented (including category "high and low"). These summaries will be presented by laboratory test and treatment group. Participants with abnormal laboratory values will be listed and values outside the normal ranges will be flagged.

The following laboratory parameters will be analyzed with respect to numerical Common Terminology Criteria for Adverse Events (CTCAE) grades, given in Table 2-3: hemoglobin, platelets, white blood cell count, neutrophils, lymphocytes, creatinine, total bilirubin (TBL), gamma-glutamyl transferase (GGT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), glucose, cholesterol, triglycerides (TG), proteinuria, electrolytes, uric acid, basophil count, hyponatremia, hypernatremia, hypokalemia, hyperkalemia, hypomagnesemia

The number and percentage of participants with CTCAE grade newly occurring or worsening after baseline will be presented. These summaries will be split into hematology and chemistry for study level reports and the pooled summary of clinical safety.

Table 2-3 CTCAE (v5.0) grades for laboratory parameters to be analyzed

CTCAE v5.0 Term	Grade 1	Grade 2	Grade 3	Grade 4
HGB decreased (Anemia)	<LLN - 10.0 g/dL; <LLN - 6.2 mmol/L; <LLN – 100 g/L	<10.0 - 8.0 g/dL; <6.2 - 4.9 mmol/L; <100 – 80 g/L	Hgb <8.0 g/dL; <4.9 mmol/L; <80 g/L; transfusion indicated	See Note 1 below
Platelet count decreased	<LLN - 75,000/mm3; <LLN – 75.0 ×10e9 /L	<75,000 - 50,000/mm3; <75.0 - 50.0 ×10e9 /L	<50,000 - 25,000/mm3; <50.0 – 25.0 ×10e9 /L	<25,000/mm3; <25.0 × 10e9 /L
White blood cell decreased	<LLN - 3000/mm3; <LLN - 3.0 × 10e9 /L	<3000 - 2000/mm3; <3.0 - 2.0 × 10e9 /L	<2000 - 1000/mm3; <2.0 - 1.0 × 10e9 /L	<1000/mm3; <1.0 × 10e9 /L
Neutrophil count decreased	<LLN - 1500/mm3; <LLN - 1.5 × 10e9 /L	<1500 - 1000/mm3; <1.5 - 1.0 × 10e9 /L	<1000 - 500/mm3; <1.0 - 0.5 × 10e9 /L	<500/mm3; <0.5 × 10e9 /L
Lymphocyte count decreased	<LLN - 800/mm3; <LLN - 0.8 × 10e9/L	<800 - 500/mm3; <0.8 - 0.5 × 10e9 /L	<500 - 200/mm3; <0.5 - 0.2 × 10e9 /L	<200/mm3; <0.2 × 10e9 /L

CTCAE v5.0 Term	Grade 1	Grade 2	Grade 3	Grade 4
Creatinine increased	>ULN - 1.5 × ULN	>1.5 - 3.0 × baseline; >1.5 - 3.0 × ULN	>3.0 × baseline; >3.0 - 6.0 × ULN	>6.0 × ULN
Proteinuria	1+ proteinuria; urinary protein <1.0 g/24 hrs	Adults: 2+ , 3+ proteinuria; urinary protein 1.0 - <3.5 g/24 hrs;	Adults: 4+ proteinuria urinary protein >=3.5 g/24 hrs;	
Hyponatremia	<LLN - 130 mmol/L	125-129 mmol/L symptomatic	120-124 mmol/L regardless of symptoms	<120 mmol/L; life-threatening consequences
Hypernatremia	>ULN - 150 mmol/L	>150 - 155 mmol/L; intervention initiated	>155 - 160 mmol/L; hospitalization indicated	>160 mmol/L; life-threatening consequences
Hypokalemia	<LLN - 3.0 mmol/L*		<3.0 - 2.5 mmol/L; hospitalization indicated	<2.5 mmol/L; life-threatening consequences
Hyperkalemia	>ULN - 5.5 mmol/L	>5.5 - 6.0 mmol/L; intervention initiated	>6.0 - 7.0 mmol/L; hospitalization indicated	>7.0 mmol/L; life-threatening consequences
Hypomagnesemia	< LLN - 0.5 mmol/L	< 0.5 - 0.4 mmol/L	< 0.4 - 0.3 mmol/L	< 0.3 mmol/L,life threatening consequencesi
Blood bilirubin increased	>ULN - 1.5 × ULN if baseline was normal; > 1.0 - 1.5 × baseline if baseline was abnormal	>1.5 - 3.0 × ULN if baseline was normal; >1.5 - 3.0 × baseline if baseline was abnormal	>3.0 - 10.0 × ULN if baseline was normal; >3.0 - 10.0 × baseline if baseline was abnormal	>10.0 × ULN if baseline was normal; >10.0 × baseline if baseline was abnormal
GGT increased	>ULN - 2.5 × ULN if baseline was normal; 2.0 - 2.5 × baseline if baseline was abnormal	>2.5 - 5.0 × ULN if baseline was normal; >2.5 - 5.0 × baseline if baseline was abnormal	>5.0 - 20.0 × ULN if baseline was normal; >5.0 - 20.0 × baseline if baseline was abnormal	>20.0 × ULN if baseline was normal; >20.0 × baseline if baseline was abnormal
ALT increased	>ULN - 3.0 × ULN if baseline was normal; 1.5 - 3.0 × baseline if baseline was abnormal	>3.0 - 5.0 × ULN if baseline was normal; >3.0 - 5.0 × baseline if baseline was abnormal	>5.0 - 20.0 × ULN if baseline was normal; >5.0 - 20.0 × baseline if baseline was abnormal	>20.0 × ULN if baseline was normal; >20.0 × baseline if baseline was abnormal
AST increased	>ULN - 3.0 × ULN if baseline was normal; 1.5 - 3.0 × baseline if baseline was abnormal	>3.0 - 5.0 × ULN if baseline was normal; >3.0 - 5.0 × baseline if baseline was abnormal	>5.0 - 20.0 × ULN if baseline was normal; >5.0 - 20.0 × baseline if baseline was abnormal	>20.0 × ULN if baseline was normal; >20.0 × baseline if baseline was abnormal
ALP increased	>ULN - 2.5 × ULN if baseline was normal; 2.0 - 2.5 × baseline if baseline was abnormal	>2.5 - 5.0 × ULN if baseline was normal; >2.5 - 5.0 × baseline if baseline was abnormal	>5.0 - 20.0 × ULN if baseline was normal; >5.0 - 20.0 × baseline if baseline was abnormal	>20.0 × ULN if baseline was normal; >20.0 × baseline if baseline was abnormal
Glucose increased (Hyperglycemia)	See Note 2 below	See Note 2 below	See Note 2 below	See Note 2 below
Glucose decreased (Hypoglycemia)	<LLN - 55 mg/dL; <LLN - 3.0 mmol/L	<55 - 40 mg/dL; <3.0 - 2.2 mmol/L	<40 - 30 mg/dL; <2.2 - 1.7 mmol/L	<30 mg/dL; <1.7 mmol/L; life-threatening consequences; seizures

CTCAE v5.0 Term	Grade 1	Grade 2	Grade 3	Grade 4
Cholesterol high	>ULN - 300 mg/dL; >ULN - 7.75 mmol/L	>300 - 400 mg/dL; >7.75 - 10.34 mmol/L	>400 - 500 mg/dL; >10.34 - 12.92 mmol/L	>500 mg/dL; >12.92 mmol/L
Hypertriglyceridemia	150 mg/dL - 300 mg/dL; 1.71 - 3.42 mmol/L	>300 mg/dL - 500 mg/dL; >3.42 - 5.7mmol/L	>500 mg/dL - 1000 mg/dL; >5.7 - 11.4 mmol/L	>1000 mg/dL; >11.4 mmol/L; life-threatening consequences

Note 1: Grade 4 Hemoglobin events are defined as life-threatening anemia events and will not be displayed in the table, as a numerical range is not provided in the CTCAE.

Note 2: Grade 1 Hyperglycemia events are defined as abnormal glucose above baseline with no medical intervention; Grade 2 Hyperglycemia events are defined as change in daily management from baseline for a diabetic, oral antiglycemic agent initiated, workup for diabetes; Grade 3 Hyperglycemia events are defined as insulin therapy initiated, hospitalization indicated; Grade 4 Hyperglycemia events are defined as life-threatening consequences. They will not be displayed in the table, as a numerical range is not provided in the CTCAE. In reporting activities, CTCAE v4.03 will be used for better descriptions.

Note 3: When grading can be either based on (i) x times LLN/ULN or (ii) x times baseline value, only grading based on x times LLN/ULN will be applied. It his the most conservative approach and can avoid different lab grading being applied at different stages of data review and reporting.

*Note 4: considering the difference between grade 1 and grade 2 hypokalemia is only an additional requirement of inventions required hence for simplicity grade 2 will not be presented separately. A manual review of grade 1 hypokalemia will be done to identify grade 2 events.

Shift tables will be presented comparing baseline laboratory result (CTCAE grade) with the worst results (expressed in CTCAE grade) during the treatment phase analyzed. Of note, baseline will be defined as last assessment prior to first dosing in core treatment phase. Participants with abnormal laboratory values will be listed and values outside the normal ranges will be flagged.

Exposure time adjusted incidence for participants with newly occurring neutropenia of CTCAE grade >= 2 will be summarized and listed in the listing.

The number and percentage of participants with clinically CTCAE grade newly occurring or worsening after baseline (treatment emergent) will be presented. Absolute and relative frequencies will be derived for non-overlap groups: CTCAE grade 1, CTCAE grade 2, CTCAE grade 3, and CTCAE grade 4.

Summaries for newly occurring or worsening clinically notable lipid abnormalities will also be provided cumulatively for each of the following parameters and categories:

- HDL:
  - $\leq$ LLN
  - $< 0.8 \times$ LLN
- LDL, cholesterol, triglycerides:
  - $\geq$ ULN
  - $> 1.5 \times$ ULN
  - $> 2.5 \times$ ULN

Participants with newly occurring or worsening after baseline abnormalities in lipid parameters will be listed. If a participant experiences newly occurring or worsening of abnormality for a parameter the entire time course of that parameter will be listed.

Newly occurring or worsening liver enzyme abnormalities will also be summarized (crude incidence as described in above) based on the event criteria given in Table 2-4:

Table 2-4 Liver-related events

Parameter	Criterion
ALT	>3×ULN; >5×ULN; >8×ULN;>10×ULN, >20×ULN
AST	>3×ULN; >5×ULN; >8×ULN >10×ULN; >20×ULN
ALT or AST	>3×ULN; >5×ULN; >8×ULN >10×ULN; >20×ULN
TBL	>1.5×ULN, >2×ULN, >3×ULN,
ALP	>2×ULN, >3×ULN. >5×ULN
ALT or AST & TBL	ALT or AST>3×ULN & TBL >2×ULN; ALT or AST >5×ULN & TBL >2×ULN; ALT or AST >8×ULN & TBL >2×ULN; ALT or AST >10×ULN & TBL >2×ULN
ALP & TBL	ALP >3×ULN & TBL >2×ULN ALP >5×ULN & TBL >2×ULN
ALT or AST & TBL & ALP	ALT or AST>3×ULN & TBL >2×ULN & ALP ≤2×ULN **(Hy's Law)** Note: elevated ALP may suggest obstruction as a consequence of gall bladder or bile duct disease; ALP may also be increased in malignancy. FDA therefore terms Hy's Law cases as indicators of *pure hepatocellular injury*. This does not mean that cases of ALT or AST >3×ULN & TBL >2×ULN & ALP >2×ULN may not result in severe DILI.

For a combined criterion to be fulfilled, all conditions have to be fulfilled on the same visit, except for Hy's law which can be fulfilled in the same period . The criteria are not mutually exclusive, e.g., a participant with ALT = 6.42×ULN is counted for ALT > 3×ULN and ALT > 5×ULN.

A table for Hy's Law (i.e., ALT or AST > 3×ULN & TBL > 2×ULN & ALP ≤ 2×ULN, and other specified events if needed) could be created. An eDISH (evaluation of Drug-Induced Serious Hepatotoxicity) plot will be provided as well.

Individual participant data listings will be provided for participants with newly occurring or worsening abnormal laboratory data. Data of participants with newly occurring or worsening liver enzyme abnormalities will be listed in an additional listing.

Boxplots over time will be presented for selected laboratory parameters (neutrophils, liver and lipid parameters).

2.7.4 Other safety data

2.7.4.1 ECG

The following quantitative variables may be summarized if requested: ventricular rate, RR interval, PR interval, QRS duration, QT interval, and corrected QT interval (QTc). Fridericia (QTcF) correction will be presented for QTc.

QTc will summarized by computing the number and percentage of participants (including 95% confidence intervals) with:

- QTc > 500 msec
- QTc > 480 msec
- QTc > 450 msec
- QTc changes from baseline > 30 msec
- QTc changes from baseline > 60 msec
- PR > 250 msec

Summary statistics will be presented for ECG variables by visit and treatment group.

A listing of all newly occurring or worsening abnormalities will be provided, as well as a by-participant listing of all quantitative ECG parameters.

2.7.4.2 Vital signs

The summary of vital signs will only include treatment emergent data, which are defined as those vital sign measurements after the first dose of study treatment and up to last dose + 180 days.

Analysis in vital sign measurement using descriptive summary statistics for the change from baseline for each post-baseline visit will be performed. These descriptive summaries will be presented by vital sign and treatment group. Change from baseline will only be summarized for participants with both baseline and post-baseline values and will be calculated as:

change from baseline = post-baseline value – baseline value

The number and percentage of participants with newly occurring notable vital signs will be presented. Criteria for notable vital sign abnormalities are provided in Table 2-5:

Table 2-5 Criteria for notable vital sign abnormalities

Vital sign (unit)	Notable abnormalities
Systolic blood pressure (mmHg)	>= 140 mmHg or < 90 mmHg
Diastolic blood pressure (mmHg)	>=90 mmHg or <60 mmHg
Pulse (bpm)	> 100 bpm or <60 bpm

2.7.4.3 Immunogenicity

2.7.4.3.1 Sample ADA Status

Each anti-drug anti-body (ADA) sample is assessed in a three-tiered ADA testing approach. All ADA samples are analyzed in the initial screening assay (first tier). Samples testing negative in the screening assay are not subject to a confirmatory assay. Samples testing positive in the screening assay are then subjected to the confirmatory assay to demonstrate that ADA are specific for the therapeutic protein product (second tier). The titer of confirmatory positive samples will be subsequently determined in the titration assay (third tier).

Samples identified as positive in the confirmatory assay are considered ADA positive and are further characterized in the neutralization assay to indicate the presence of neutralizing antibodies (NAb). In addition, samples identified as positive in the confirmatory assay will be

tested for cross- reactivity with the endogenous protein in confirmatory assay with ANGPTL3. If antibodies cross reacting with ANGPTL3 are detected, the samples will be further tested for cross reactivity to ANGPTL4.

Samples can test negative in either the screening or confirmatory assay but for statistical analysis purposes they are not differentiated. The following properties of each sample will be provided in the source data (i.e., the third party data output (e.g., WLIMS) processed by PreAdvance):

- Result of assay according to pre-specified confirmatory cut point: 'POSITIVE', 'NEGATIVE', or 'NOT REPORTABLE'
- Titer: numerical representation of the magnitude of ADA response
- Presence of NAb (for positive samples, if NAb assay results are available): 'POSITIVE' or 'NEGATIVE'
- Threshold for determining treatment-boosted (titer fold change (i.e., 4-fold))

The following definitions apply only to non-missing samples:

- ADA-negative sample: Sample where assay result is 'NEGATIVE'
- ADA-positive sample: Sample where assay result is 'POSITIVE'
- ADA-positive NAb sample: ADA-positive sample where NAb assay result is 'POSITIVE'
- ADA-negative NAb sample: ADA-positive sample where NAb assay result is 'NEGATIVE'
- ANGPTL3-crossreactive sample: ADA-positive sample where confirmatory assay result with ANGPTL3 is 'POSITIVE'
- ANGPTL4-crossreactive sample: ADA-positive sample where confirmatory assay result with ANGPTL4 is 'POSITIVE'

The following definitions apply only to post-baseline ADA-positive samples with a corresponding non-missing baseline sample. To be classified as treatment-induced, treatment-boosted or treatment-unaffected, both the post-baseline and baseline titer must be non-missing.

- treatment-induced ADA-positive sample: ADA-positive sample post-baseline with ADA-negative sample at baseline
- treatment-boosted ADA-positive sample: ADA-positive sample post-baseline with titer that is at least the titer fold (i.e., 4-fold) change greater than the ADA-positive baseline titer
- treatment-unaffected ADA-positive sample: ADA-positive sample post-baseline with titer that is less than the titer fold (i.e., 4-fold) change greater than the ADA-positive baseline titer

The following summaries of ADA sample status (n and %) will be provided using the *Immunogenicity prevalence set*:

- ADA-positive samples (ADA prevalence), ADA-positive NAb samples, ANGPTL3-crossreactive samples, and ANGPTL4-crossreactive samples, both overall and by time point (including baseline). For summaries by time point, the denominator is the number of subjects at that time point with a non-missing sample.

Listings will be provided of ADA sample status together with the corresponding PK concentration (if available).

2.7.4.3.2 Subject ADA status

Subject ADA status is defined as follows:

- *Treatment-induced ADA-positive subject*: subject with ADA-negative sample at baseline and at least one treatment-induced ADA-positive sample
- *Treatment-boosted ADA-positive subject*: subject with ADA-positive sample at baseline and at least one treatment-boosted ADA-positive sample
- *Treatment-unaffected ADA-positive subject*: subject with ADA-positive sample at baseline, no treatment-boosted ADA-positive samples, and at least one treatment-unaffected ADA-positive sample
- *Treatment-reduced ADA-positive subject*: subject with ADA-positive sample at baseline and at least one non-missing post baseline sample, all of which are ADA-negative samples
- *ADA-negative subject*: subject with ADA-negative sample at baseline and at least one non-missing post baseline sample, all of which are ADA-negative samples
- *Inconclusive subject*: subject who does not qualify for any of the above definitions or a subject for which the baseline sample is missing

The following summaries of ADA subject status (n and %) will be provided using the *Immunogenicity incidence set* (for % the denominator is the number of subjects in the *Immunogenicity incidence set* unless otherwise specified):

- Subjects with ADA-negative sample at baseline
- Subjects with ADA-positive sample at baseline
- Subjects with ADA-positive NAb sample at baseline
- Subjects with ANGPTL3-crossreactive sample
- Subjects with ANGPTL4-crossreactive sample
- ADA-negative subjects
- Treatment-induced ADA-positive subjects; for % the denominator is the number of subjects with ADA-negative sample at baseline.
- Treatment-boosted ADA-positive subjects; for % the denominator is the number of subjects with ADA-positive sample at baseline.
- ADA-inconclusive subjects
- ADA-positive subjects (i.e. ADA incidence): calculated as the number of treatment-boosted ADA-positive and treatment-induced ADA-positive subjects.

- ADA-positive NAb subjects: calculated as the number of ADA-positive subjects with at least one NAb sample (baseline or post-baseline); for % the denominator is the number of ADA-positive subjects.

Listings will be provided of ADA subject status.

2.7.4.3.3 Transient and Persistent ADA responses

Treatment-induced ADA-positive subjects (if any) will be analyzed for transient and persistent ADA responses. Summaries will be provided to present % of transient and persistent responses.

Transient ADA response is defined as follows:

- Treatment-induced ADA detected only at one sampling time point during the treatment or follow-up observation period (excluding the last sampling time point) and that sampling time point is 16 weeks or more before an ADA-negative last sample

OR

- Treatment-induced ADA detected at two or more sampling time points during the treatment (including follow-up period if any), where the first and last ADA-positive samples (irrespective of any negative samples in between) are separated by a period less than 16 weeks, and the subject's last sampling time point is ADA-negative

Persistent ADA response is defined as follows:

- Treatment-induced ADA detected at two or more sampling time points during the treatment (including follow-up period if any), where the first and last ADA-positive samples (irrespective of any negative samples in between) are separated by a period of 16 weeks or longer

OR

Treatment-induced ADA incidence at the last sampling time point of the treatment or follow-up observation period, or at a sampling time point within less than 16 weeks before an ADA-negative last sample

2.7.4.4 Acute Inflammatory Reactions (AIR)

Incidence of acute inflammatory reactions (see Section 2.10 for details) will be summarized by treatment group and visit separately from summaries of AEs.

2.8 Pharmacokinetic endpoints

LNA043 and ANGPTL3 serum concentration data will be listed by treatment, participant, and visit/sampling time point.

Descriptive summary statistics will be provided by treatment and visit/sampling time point, including the frequency (n, %) of concentrations below the Lower Limit of Quantification (LLOQ) and reported as zero.

All completed participants with quantifiable pharmacokinetic (PK) measurements and who received at least 80% of the planned dose on the day of PK assessment will be included in the PK data analysis. Serum concentrations will be expressed in mass per volume units (ng/mL). In addition to mean, standard deviation (SD), coefficient of variation (CV), median and quartiles, the geometric mean and geometric coefficient of variation (CV), minimum and maximum will be presented. When the dataset includes zero values the geometric mean and geometric coefficient of variation (CV) will be calculated for the subset of non-zero values; n the number of included participants and n(log) the number of included participants with non-zero values will be presented. The formula for deriving the geometric mean and CV (%) is as follows:

- CV (%) = (SD/mean)*100
- geometric mean = exp((sum of log transformed data) / number of non-missing data points after log transformation)
- geometric CV = sqrt(exp(variance of log-transformed data)-1)*100

Concentrations below LLOQ will be imputed by 0 in summary statistics. No pharmacokinetic parameters will be calculated in this study as sparse sampling does not allow valid estimation by non-compartmental analysis.

In addition, sample number, concentration, sample date, sample time and elapsed time from first dose (in days) will be listed by treatment sequence.

Modeling of the data may be performed as appropriate. During modeling of the pharmacokinetics of the study drugs, the broad principles outlined in [FDA Guidance for Industry: Population Pharmacokinetics] will be followed. As the PK data from the current study may be pooled with data from previous studies, the PK modeling will be described and reported separately.

2.9 PD and PK/PD analyses

Exploratory analysis to investigate the correlation between the PK data and efficacy outcomes may be performed.

Results will be reported separately.

2.10 Patient-reported outcomes

Patient reported outcomes will be evaluated based on FAS (except for AIR, which will be analyzed on SAF) unless otherwise specified.

Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)

The WOMAC has been developed to assess the course of disease or response to treatment in participants with knee or hip osteoarthritis (Collins et al 2011). It includes 24 items and three subscales:

1. pain severity during various positions or movements ("WOMAC Pain"): 5 items
2. severity of joint stiffness ("WOMAC Stiffness"): 2 items
3. difficulty performing daily functional activities ("WOMAC Function"): 17 items

The 24 items will be scored with a numeric rating scale using 11-box horizontal scale, with the left end (0) marked as "none" and the right end (10) marked as "extreme". The recall period is 48 hours.

If 2 or more pain items, both stiffness items, and 4 or more physical function items are missing, the response should be regarded as invalid and the deficient subscale(s) should not be used in analysis.

Higher scores indicate worse pain, stiffness, or physical function. WOMAC score will be rescaled to a normalized 0-100 scale in the analysis:

WOMAC score (0-100 scale) = sum of non-missing scores / number of non-missing item × 10

For WOMAC pain and WOMAC function (see Section 2.6 for details) Summary statistics of observed data will be provided for change from baseline in WOMAC total score by visit and treatment. Total WOMAC will be defined as the sum of the sub-scale scores (pain, stiffness, function. If any of the sub-scale scores are missing, then the total score will be considered missing.

Short Form 12 Health Survey (SF-12)

The SF-12 is one of the most commonly used generic health-related quality of life (HRQL) measures in participants with hip or knee disorders. The SF-12 has a lower responder burden compared to SF-36 and has been validated for use in knee OA trials (Gandhi et al 2001). The SF-12 contains a subset of 12 items from the SF-36, including one or two items from each of the eight SF-36 scales (Gandek et al 1998). Two items are included from the Physical Functioning and Mental Health scales because these scales have been shown to best predict physical and mental health; two items each are also included from both Role Functioning scales, because these are relatively coarse scales. One item each is included from the remaining four scales. Information from all 12 items is used to construct physical and mental component summary measures.

The following variables will be evaluated:

- SF-12 physical component summary
- SF-12 mental component summary

Summary statistics of observed data will be provided for change from baseline by visit and treatment.

Patient's Global Assessment of Pain (PGA)

The PGA is commonly used in OA trials (McAlindon et al 2015). Participants will answer the question: "Considering all the ways your osteoarthritis of the knee has affected you during the last 48 hours, select the number that best describes the impact of your knee osteoarthritis on your daily life?" by using an 11-box horizontal scale, with the left end marked as “none” and the right end marked as “extreme”.

Summary statistics of observed data will be provided for change from baseline in PGA by visit and treatment.

Patient Global Impression of Severity (PGIS) and Patient Global Impression of Change (PGIC)

The PGIS and PGIG provide a responsive and readily interpretable measure of participant's assessments of the clinical importance of their improvement or worsening of their knee pain during functional performance outcome assessments over the course of the clinical trial (Dworkin et al 2005). Participants will answer questions about their perception of disease activity by using a 4-point and a 5-point scale, for PGIS and PGIC, respectively.

Summary statistics of observed data will be provided for change from baseline in PGIS by visit and treatment.

A shift table will be provided for PGIC by visit and treatment.

Widespread Pain Index (WPI) and Symptom Severity Scale (SS scale)

WPI and SS scale will be collected only at Screening Visit 2 to exclude participants with fibromyalgia, according to recent criteria by the American College of Rheumatology (Wolfe et al 2010). The use of both scales allows for highly accurate identification and differentiation of fibromyalgia participants.

The WPI is a 19-point checklist that assesses the presence of pain or tenderness (within the prior seven days) in 19 specific areas of the body; each affected area receives one point. The WPI score is between 0 and 19.

The SS scale score is the sum of the severity of the 3 symptoms (fatigue, waking unrefreshed, cognitive symptoms) plus the extent (severity) of somatic symptoms in general. The final score is between 0 and 12.Summary statistics of observed data will be provided for WPI and SS scale at baseline by treatment.

Acute Inflammatory Reactions (AIR)

Acute inflammatory reactions (AIRs), are defined as increase of pain by 3 points on an 11-points NRS (range 0-10) and a self-reported synovial fluid effusion (i.e., joint swelling) within 3 days following i.a. injection documented in the participant eDiary.

Synovial fluid effusions during the treatment period are reported in the participant eDiary by answering 'Yes' to swelling. If it occurs within 3 days after i.a. injection, then the participant is classified as having a self-reported synovial fluid effusion within 3 days following i.a. injection. Given the injection happens on Day 0, the 'within 3 days' is defined as Day 1, Day 2 and Day 3.

Participants are classified as experiencing an increase of pain by 3 points on an 11-points NRS from pre-dose to post-dose when the following occurs within 3 days following i.a. injection:

CCI [redacted]

Incidence of acute inflammatory reactions (AIR) will be summarized by treatment group and visit.

2.11 Biomarkers

Results for soluble biomarkers will be reported separately.

2.12 Pharmacogenetics

Exploratory pharmacogenetics studies are designed to investigate the association between genetic factors (genotypes) and clinical assessments (phenotypes) which are collected during the clinical trial.

The study includes an optional genetic research component which requires a separate informed consent signature if the participant agrees to participate. The purpose of genetic research may be to better understand the safety and efficacy of LNA043, or to learn more about human diseases, or to help develop ways to detect, monitor and treat diseases.

Without prior evidence of a strong association, a number of possible associations will be evaluated with exploratory analyses. A range of statistical tests (chi-square tests, analysis of covariance (ANCOVA), linear and logistic regression) will be used for the analyses. Additional data, from other clinical trials, are often needed to confirm associations. Alternatively, if the number of available samples is too small to complete proper statistical analyses, the data may be combined, as appropriate, with those from other studies to enlarge the dataset for analysis.

Data generated on hypothesis-free platforms may be reported separately (e.g., CSR addendum).

2.13 Other Exploratory analyses

The following exploratory efficacy variables are planned to be analyzed using the FAS.

1. Change from baseline in cartilage thickness in the cMTFC assessed by qMRI of the target knee at Weeks CCI [redacted]
2. Change from baseline at Weeks CCI [redacted] in:

- WOMAC pain (on a normalized 0-100 scale)
- WOMAC pain walking on flat surface response (on a normalized 0-100 scale)
- WOMAC function (on a normalized 0-100 scale)

3. Change from baseline in cartilage thickness in the total, medial and lateral TFCs assessed by qMRI at Weeks CCI
4. Proportion of participants demonstrating at Week CCI structural progression in the target knee defined as
   - Loss of medial minJSW ≥ 0.70 mm from baseline by X-ray
5. Change from baseline in physical function at Weeks CCI
   - 40-meter (4×10m) fast-paced walk test
   - 30-second chair stand test
   - 6-minute walking test
6. Change from baseline in individual WOMAC pain questions (Walking on flat surfaces, Walking up or down stairs, While sitting or lying down, While standing (on a normalized 0-100 scale) at Weeks CCI
7. Change from baseline in WOMAC stiffness score (on a normalized 0-100 scale) at Weeks CCI
8. Change from baseline in WOMAC total score (on a normalized 0-300 scale) at Weeks CCI
9. Change from baseline at Weeks CCI in OA disease activity as assessed by the PGA, PGIS and PGIC
10. Change from baseline at Weeks CCI in QoL using:
    - SF-12 physical component summary
    - SF-12 mental component summary
11. CCI at Weeks CCI
    - CCI
      s.
      - CCI

12. Change from baseline at Weeks CCI in the following assessments
    - CCI
    - CCI
    - CCI
    - CCI
        - CCI
13. Change from baseline to Weeks CCI
    - Fixed-location medial JSW measured on X-ray
    - Medial minJSW measured on X-ray
14. Change from baseline to Week 104 in OA pain in the target knee as assessed by the CCI
15. Proportion of participants who achieved WOMAC pain improvement by >=10 points on a normalized 0-100 scale
16. Proportion of participants who achieved WOMAC function improvement >=10 points on a normalized 0-100 scale
17. Proportion of OMERACT-OARSI responders at Weeks CCI
18. Change from baseline to Week 104 in actigraphy-derived metrics including: Minutes of moderate to vigorous activity per day, Total steps per day, Average duration of walking bouts (>=2min) within a day, Average walking cadence (for bouts >=2min) per day, and 90th percentile of cadence, a measure of peak performance, representing the fastest cadence over the measurement period (per day).
19. Change from baseline in CCI during the study

Analyses may include summary statistics and listings and as appropriate graphical presentations of the original values and the changes from baseline.

Additional exploratory analyses may be reported separately, outside of the CSR and will be defined in the Data Exploration Strategy Document.

2.14 CCI

CCI

2.15 Interim analysis

The primary efficacy analysis will be performed at Week 104 to assess efficacy and safety variables. The final analysis at Week 260 will be performed as a safety and efficacy follow-up assessment. CCI

When an interim analysis is performed, designated Novartis team members will be unblinded, however core Novartis team reviewing study data will remain blinded until the primary efficacy (Week 104) database lock.

The statistical methodology and other details regarding the interim analyses will be defined in a separate interim analysis SAP document.

3 Sample size calculation

3.1 Primary endpoint(s)

A total of CCI participants per each group is deemed appropriate to achieve adequate power for the primary and key secondary endpoints for this study.

The mean change from baseline in cartilage thickness of the cMTFC for placebo at Week 52 is -0.116 mm with the standard deviation of CCI for placebo based on the results of the Galapagos and Servier ROCCELLA Phase 2 study, which enrolled a comparable knee OA population (Bernard et al 2020). Cartilage loss is assumed to be constant over a 2-year period as shown in the sprifermin trial (Hochberg et al 2019), therefore a -0.23 mm mean change from baseline in cartilage thickness of the cMTFC for placebo at Week 104 is used for this study. Assuming that at the very least LNA043 stops further cartilage degradation compared to placebo, a treatment difference of 0.23 mm in favor of LNA043 with a common standard deviation of CCI will be observed at Week 104, a sample size of CCI participants in each treatment group yields approximately CCI power at $\alpha = 0.05$ (one-sided) using Dunnett contrasts (Addplan DF 4.0.9).

3.2 Secondary endpoint(s)

As it is also important to show an effect on a clinical endpoint, a sample size calculation was done using the WOMAC pain subscale. It was shown in the EUFLEXXA trial (Altman et al 2009) that a change from baseline at Week 26 using intra-articular saline demonstrated a mean of 16.3 mm with standard deviation of 26.8 on the WOMAC pain subscale. A treatment difference of CCI in favor of LNA043 at Week 104 will need a sample of CCI power at $\alpha = 0.05$ (one-sided) using Dunnett contrasts (Addplan DF 4.0.9). Considering a CCI dropout rate, a sample size of CCI participants in each treatment group will be needed.

4 Change to protocol specified analyses

- Only physical component and mental component score, but no total score of SF-12v2 will be presented
- Use of prohibited medications is considered an intercurrent event for the primary and secondary endpoints
- The sensitivity analysis for the primary endpoint will use a MMRM instead of an ANCOVA model.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

Any partial dates will be imputed as follows:

We take the earlier day of

- The last day in the month and
- The end day of the corresponding epoch

5.1.2 AE date imputation

Impute AE end date:

1. If the AE end date 'month' is missing, the imputed end date should be set to the earliest of the (min (last visit date, last dose date + 180 days), 31DECYYYY, date of death).
2. If the AE end date 'day' is missing, the imputed end date should be set to the earliest of the (min (last visit date, last dose date + 180 days), last day of the month, date of death).
3. If AE 'year' is missing or AE is ongoing, the end date will not be imputed.

Impute AE start date:

Before imputing AE start date, find the AE start reference date as below

- If the (imputed) AE end date is complete and the (imputed) AE end date < treatment start date then AE start reference date = min(informed consent date, earliest visit date).
- Else AE start reference date = treatment start date

1. If the AE start date 'year' value is missing, the date uncertainty is too high to impute a rational date. Therefore, if the AE year value is missing, the imputed AE start date is set to NULL.
2. If the AE start date 'year' value is less than the treatment start date year value, the AE started before treatment. Therefore:
    a. If AE 'month' is missing, the imputed AE start date is set to the mid-year point (01JulYYYY).
    b. Else if AE 'month' is not missing, the imputed AE start date is set to the mid-month point (15MONYYYY).
3. If the AE start date year value is greater than the treatment start date year value, the AE started after treatment. Therefore:
    a. If the AE month is missing, the imputed AE start date is set to the year start point (01JanYYYY).
    b. Else if the AE month is not missing, the imputed AE start date is set to the later of (month start point (01MONYYYY), AE start reference date + 1 day).
4. If the AE start date year value is equal to the treatment start date year value:
    a. And the AE month is missing the imputed AE start date is set to the AE reference start date + 1 day.
    b. Else if the AE month is less than the treatment start month, the imputed AE start date is set to the mid-month point (15MONYYYY).
    c. Else if the AE month is equal to the treatment start date month or greater than the treatment start date month, the imputed AE start date is set to the later of (month start point (01MONYYYY), AE start reference date + 1 day).

If complete (imputed) AE end date is available and the imputed AE start date is greater than the (imputed) AE end date, then imputed AE start date should be set to the (imputed) AE end date.

5.1.3 Concomitant medication date imputation

Impute concomitant medication (CM) end date:

1. If CM end day is missing and CM month/year are non-missing then impute CM day as the minimum of treatment end date and the last day of the month.
2. If CM end day/month are missing and CM year is non-missing then impute CM day as the minimum of treatment end date and the end of the year (31DECYYYY).
3. If imputed CM end date is less than the CM start date, use the CM start date as the imputed CM end date.

Impute CM start date:

1. If the CM start date year value is missing, the imputed CM start date is set to one day prior to treatment start date.

2. If the CM start date year value is less than the treatment start date year value, the CM started before treatment. Therefore:
   a. If the CM month is missing, the imputed CM start date is set to the mid-year point (01JulYYYY).
   b. Else if the CM month is not missing, the imputed CM start date is set to the mid-month point (15MONYYYY).
3. If the CM start date year value is greater than the treatment start date year value, the CM started after treatment. Therefore:
   a. If the CM month is missing, the imputed CM start date is set to the year start point (01JanYYYY).
   b. Else if the CM month is not missing, the imputed CM start date is set to the month start point (01MONYYYY).
4. If the CM start date year value is equal to the treatment start date year value:
   a. And the CM month is missing or the CM month is equal to the treatment start date month, then the imputed CM start date is set to one day prior treatment start date.
   b. Else if the CM month is less than the treatment start date month, the imputed CM start date is set to the mid-month point (15MONYYYY).
   c. Else if the CM month is greater than the treatment start date month, the imputed CM start date is set to the month start point (01MONYYYY).

If complete (imputed) CM end date is available and the imputed CM start date is greater than the (imputed) CM end date, then imputed CM start date should be set to the (imputed) CM end date.

5.1.3.1 Prior therapies date imputation

See Section 5.1.3.

5.1.3.2 Post therapies date imputation

See Section 5.1.3.

5.1.4 First diagnosis date imputation

1. If the first diagnosis day/ month are missing and the year is non-missing:
   a. If the year part of the first diagnosis date is equal to the year part of the inform consent date, then the imputed first diagnosis date is set to the year start point (01JanYYYY).
   b. Otherwise the imputed first diagnosis date is set to the mid-year point (01JulYYYY).
2. If the first diagnosis day is missing and the month/year are non-missing:
   a. If the month and year part of the first diagnosis date is equal to the month and year part of the inform consent date, then the imputed first diagnosis date is set to the month start point (01MONYYYY).
   b. Otherwise the imputed first diagnosis date is set to the mid-month point (15MONYYYY).

5.1.5 Other imputations

5.1.5.1 Laboratory data

For laboratory test values below Lower Level of Quantification (LLQ) or above Upper Level of Quantification (ULQ) will be imputed as LLQ or ULQ value, respectively. The numerical part of the reported result will be treated as the actual LLQ or ULQ. These laboratory values will be displayed in listings using the standard unit with the reported sign ("<" or ">").

5.1.5.2 Performance based tests

The performance based tests include: 30 second chair stand test, 4×10 m fast-paced walk test, and 6 minutes walking distance test.

The adaptation is considered as follows:

- 30 second chair stand test: if the participant cannot stand even once then allow the hands to be placed on the legs or use the regular mobility aid.
- 4×10 m fast-paced walk test: if the participant uses regular walking aid
- 6 minutes walking distance test: if the participant uses regular walking aid

This is then scored as an adapted test score. In this case, the adapted score, instead of the actual score, and the adaptations will be collected in the CRF. In case of adaptation, the actual score will be imputed as 0 and used in the analysis as stated in Section 2.6. Additional analyses (i.e., summary tables, listings) of adaptation and/or adapted score may be provided per request.

5.2 AEs coding/grading

Adverse events will also be coded according to MedDRA dictionary, using a narrow search. The MedDRA version used for reporting the adverse events will be described in a footnote. Safety topics of interest, such as risks defined in the Safety Profiling Plan, topics of interest regarding signal detection or routine analysis are defined in the Program Case Retrieval Sheet.

5.3 Laboratory parameters derivations

See Table 2-3 and Table 2-4.

5.4 CCI


CCI

5.5 Statistical models

5.5.1 Analysis supporting primary objective(s)

5.5.1.1 Summary statistics for continuous data

Summary statistics (including N, mean, standard deviation, minimum, lower quartile, median, upper quartile and maximum) will be provided for continuous data by visit and treatment group.

5.5.1.2 Analysis of covariance

Univariate model

An analysis of covariance (ANCOVA) model will be used to analyze some endpoints. The model will include factors and covariates as specified for respective analysis. The SAS code below outlines a template for the analysis where covariates can be added or removed as required:

ods output diffs=*lsm_diff* lsmeans=*lsm_est*;

proc mixed data=*mydata*;

by *visit*;

class *treatment*;

model *outcome* = *treatment baseline_score*;

lsmeans *treatment* / diff cl;

run;

Least-square-mean (LSM) estimates for each treatment group and LSM difference, confidence intervals and p-value for the difference between LNA043 and placebo can be obtained.

Repeated measures analysis

Some endpoints will be analyzed using a longitudinal model that comprises several visits. The model used will be mixed model repeated measures (MMRM) with factors, covariates, interactions and covariance structure as specified for respective analysis. The SAS code below outlines a template for the analysis where covariates and interaction terms can be added or removed as required:

ods output diffs=*lsm_diff* lsmeans=*lsm_est*;

proc mixed data=*mydata*;

class *treatment visit*;

model *outcome = treatment visit baseline_score treatment*visit*

*baseline_score *visit* / ddfm=kr;

repeated *visit* / type=un subject=;

lsmeans *treatment*visit* / diff cl adjust=dunnett;

run;

Least-square-mean (LSM) estimates for each treatment group and LSM difference, confidence intervals and p-value for the difference between LNA043 and placebo will be calculated at appropriate analysis visits.

In case the MMRM model does not converge the following sequential steps will be used:

1. change to type=cs. If still no convergence, perform step 2.
2. remove covariate *baseline_score*visit*.

For the selection of appropriate one-sided p-value use the following logic:

For endpoints where superiority is defined as LNA less than placebo:
If LNA is considered superior when its mean is less than the mean of placebo (as defined in the SAP) and the ESTIMATE (from the data) is less than 0, then use ADJP/2.

If LNA is considered superior when its mean is less than the mean of placebo (as defined in the SAP) and the ESTIMATE (from the data) is greater than or equal to 0, then use 1-ADJP/2.

For endpoints where superiority is defined as LNA greater than placebo:

If LNA is considered superior when its mean is greater than the mean of placebo (as defined in the SAP) and the ESTIMATE (from the data) is greater than 0, then use ADJP/2.

If LNA is considered superior when its mean is greater than the mean of placebo (as defined in the SAP) and the ESTIMATE (from the data) is less than or equal to 0, then use 1-ADJP/2.

5.5.2 Analysis supporting secondary objective(s)

5.5.2.1 Summary statistics for binary and categorical data

Summary statistics for discrete variables will be presented in contingency tables and will include count and frequency in each category. If applicable, confidence intervals will be derived as well based on the score method including continuity correction (Newcombe 1998):

With z as $(1-\alpha/2)$-quantile of the standard normal distribution (SAS: z = probit(1-α/2), n as total number of participants (i.e., number of participants in the denominator), p as estimated crude incidence (number of participants with event / n) and $q = 1-p$

Then the lower limit is

$$L = 100 \times \max\left(0, \frac{2np + z^2 - 1 - z\sqrt{z^2 - 2 - 1/n + 4p(nq+1)}}{2(n+z^2)}\right)$$

and the upper limit is

$$U = 100 \times \min\left(1, \frac{2np + z^2 + 1 + z\sqrt{z^2 + 2 - 1/n + 4p(nq-1)}}{2(n+z^2)}\right).$$

In addition, if L > p then L = p and if U < p then U = p.

For binary response variables the placebo-adjusted response rates including 95% confidence interval will be derived.

SAS code for risk difference:

```
proc freq data=;
    tables response* treatment / riskdiff;
run;
```

Note the response value should be sorted with '1' ahead of '0'.

5.5.2.2 Logistic regression

Certain binary outcome variables, e.g., response outcomes, will be evaluated using a logistic regression model with treatment as factor and baseline score (if applicable) as a covariate. The marginal standardization method will be used to calculate the mean response rate in each treatment group as well as their difference. This method uses the same fitted logistic model, but

involves using the model to predict, for each participant in the study, the mean outcome assuming assignment to each particular treatment group in turn, assuming each participant's observed values for the other baseline covariates (i.e., baseline score). Averaging these predictions for each treatment group provides the estimate of the mean response rate for each treatment group. Then the difference will be derived based on the estimated mean response rates comparing LNA043 regimen versus placebo.

The macro Margins ([Predictive margins and average marginal effects]) will be used.

SAS code example as the following,

%Margins(data = *mydata*,

class = *treatment*,

response = *response*,

roptions = event='1',

model = *treatment baseline_score*,

dist = binomial,

margins = *treatment*,

options = cl diff)

For cases where the convergence status indicates that the model did not reach appropriate convergence (conv_status is not 0), no risk difference or p-value will be presented from that model.

However, if the issue relates to the primary timepoint of Week 104 then the following steps will be followed:

1. Remove *baseline_score* from the model. If there are still issues perform step 2
2. Use Fisher's exact test as described below.

When Fisher's exact test is applied, only a p-value for a test of equal response in the two groups can be obtained (no risk difference or confidence intervals can be estimated.)

ods output fishersexact=*fisher*;

proc freq data=*mydata*;

by *visit*;

table *treatment***response* / fisher;

run;

Input dataset should only contain data from the two treatment groups to be compared.

5.5.2.3 Multiple Imputation

A linear regression model will be used to perform multiple imputation (MI) under a missing-at-random (MAR) assumption. To help preserve the relationship between outcome and covariates within each treatment a separate model will be run for each treatment. This will also

help ensure that the imputation model does not make stronger assumptions on data relations than the analysis model.

For logistic regression models binary outcomes are derived from continuous variables. Multiple imputation will be performed on the underlying continuous variable directly.

The SAS code below outlines a template for the analysis where covariates and visits can be added or removed as required. To ensure that results can be replicated the data should be sorted by participant number before running the model (the data should be in horizontal format with one participant per dataset row).

```
proc mi data=mydata seed=12202 minimum=&minval maximum=&maxval
minmaxiter=10000000 nimpute=100 out=mi_out;
    by treatment;
    fcs reg (/details);
    var value1 value2 value3;
run;
```

Where in the template code the continuous variable to be imputed is *value* (e.g., *value1* could be the baseline value and *value2* the first post-treatment measurement of the variable to be imputed.) Normally, all data collection visits during the analysis period of interest would be included in the model. Including variables using a CLASS statement instead of a BY statement should help facilitate model convergence also when the number of non-missing data points are low for some specific covariate level and visit combination. The FCS option is used to ensure that also non-monotone missing data can be handled in an appropriate way.

For a situation where several variables need to be imputed using separate models (e.g., using independent models to impute each component needed to derive a response variable *V*) a step-wise process needs to be implemented as outlined below:

1. Run the SAS code as described above for the first variable to be imputed
2. Run the SAS code as described above for the next variable to be imputed (but with the following changes: "data=*mi_out*", "out=*mi_out2*", "by *treatment _imputation_*", "nimpute=1")
3. Repeat step 2, but with input dataset equal to the output dataset from the prior step, until all *j* variables have been imputed resulting in a dataset named mi_out_j
4. Derive the variable *V* from within mi_out_j

The required analysis (e.g., MMRM, logistic regression) is then performed separately within each imputation dataset (as identified by variable *_imputation_*). To obtain the final result of the imputation process the analysis result from each imputation dataset needs to be combined according to Rubin's rules as outlined below:

```
ods output parameterestimates=mi_result;
proc mianalyze data=;
    modeleffects estimate;
```

stderr *estimate_se*;

run;

The *estimate* and *estimate_se* parameters come from the analysis model used to analyze the imputed variable within each imputation dataset (e.g., from the lsmean estimate of the treatment difference and its standard error obtained from PROC LOGISTIC or PROC MIXED.)

To obtain binary response rates and confidence intervals for individual treatment groups the following process should be followed (exemplified for one visit):

ods output binomialprop=*bin_est*;

proc freq data=;

by *treatment _imputation_*;

table *response* / binomial (cl=wilson correct);

run;

Then apply a logit transformation on the saved proportions and derive its standard error:

data *bin_est*;

set *bin_est*;

estimate=log(*_bin_*/(1-*_bin_*));

estimate_se=*e_bin*/(*_bin_**(1-*_bin_*));

run;

The transformed binomial proportion estimates and its standard errors are then combined by applying Rubin's rules as described above using PROC MIANALYZE. Before presenting the combined data it needs to be transformed back as follows:

data *mi_result*;

set *mi_result*;

prop_est=1/(1+exp(-*estimate*));

prop_lower=1/(1+exp(-*lclmean*));

prop_upper=1/(1+exp(-*uclmean*));

run;

If all responses are imputed the same as 0 (or 1) for all imputation datasets for a specific treatment or subgroup then the between-imputation-variation will be zero. The combined final response rate for the specific treatment or subgroup would be presented as seen in any of the imputed datasets together with 95% CI from Wilson's method (as obtained from PROC FREQ).

The following steps will be performed to handle special cases:

- If after imputation all responses (observed+imputed) are the same either 0 or 1 for all imputation datasets for a specific treatment or subgroup it will not be possible to perform a

logit transformation and the response rate (0% or 100%) for these cases will be presented together with the 95% CI from Wilson's method (as obtained from PROC FREQ).

- If after imputation the average response rate is the same across all imputed datasets (but not 0 or 1) there is no between dataset variation and Rubin's rules cannot be applied. Instead the average response will be used with 95% CI from Wilson's method (as obtained from PROC FREQ)

5.5.3 Crude incidence and related risk estimates

5.5.3.1 Crude incidence and 100*(1-α)% confidence interval

For n participants, each at risk to experience a certain event with probability π, the crude incidence is estimated as $p=x/n$, where x is the number of participants with the event.

Absolute and relative frequencies will be displayed as well as 95% confidence interval for the relative frequency based on the score method including continuity correction (Newcombe 1998).

With z as $(1-\alpha/2)$-quantile of the standard normal distribution (SAS: z = PROBIT(1-α/2), n as total number of participants (i.e., number of participants in the denominator), and p as estimated crude incidence (number of participants with event / n) it is $q=1-p$.

Then the lower limit is

$$L = \max\left(0, \frac{2np + z^2 - 1 - z\sqrt{z^2 - 2 - 1/n + 4p(nq+1)}}{2(n+z^2)}\right)$$

and the upper limit is

$$U = \min\left(1, \frac{2np + z^2 + 1 + z\sqrt{z^2 + 2 - 1/n + 4p(nq-1)}}{2(n+z^2)}\right).$$

In addition, if L > p then L = p and if U < p then U = p.

If appropriate, an exact 100*(1-α)% confidence interval (Clopper-Pearson, 1934) will be obtained by using the SAS procedure PROC FREQ with the EXACT BINOMIAL statement. However, the confidence interval derived via the score method including continuity correction will be the default in safety analyses.

5.5.3.2 Odds ratio and 100*(1-α)% confidence interval

For an investigational drug group with n_1 participants at risk, independent from the control group (e.g., placebo or comparator) with n_0 participants at risk, of whom x_1 and x_0 experience a certain event with probability π_1 and π_0 respectively, the odds ratio is estimated as

$\frac{p_1/(1-p_1)}{p_0/(1-p_0)}$ with $p_1 = x_1/n_1$ and $p_0 = x_0/n_0$. A conditional exact 100*(1-α)% confidence interval can be obtained by using the SAS procedure PROC FREQ with statement EXACT OR. However, to be able to adjust for covariates odds ratios will primarily be obtained from PROC LOGISTIC.

5.5.3.3 Risk difference and 100*(1-α)% confidence interval

For an investigational drug group with n_1 participants at risk, independent from the control group (e.g., placebo or comparator) with n_0 participants at risk, of whom x_1 and x_0 experience a certain event, the risk difference is estimated as p_1-p_0 with $p_1 = x_1/n_1$ and $p_0 = x_0/n_0$.

Exact unconditional confidence limits for the risk difference can be obtained with SAS procedure PROC FREQ and option RISKDIFF in the TABLES statement, specifying the RISKDIFF option also in the EXACT statement.

5.5.4 Exposure adjusted incidence rate and related risk estimates

5.5.4.1 Exposure adjusted incidence rate and 100*(1-α)% confidence interval

It will be assumed that for each of n participants in a clinical trial the time t_j (j=1,...,n) to the first occurrence of a certain event is observed, or if the event was not experienced, the (censored) time to the end of the observation period. The sequence of first occurrences of an event will be modeled to follow approximately a Poisson process with constant intensity θ. The rate parameter θ will be estimated as $\lambda = D/T$, where $T = \sum_{j=1}^{n} t_j$ and D is the number of participants with at least one event. Conditionally on T, an exact 100*(1-α)% confidence interval for a Poisson variable with parameter θT and observed value D can be obtained based on (Garwood 1936), from which an exact 100*(1-α)% confidence interval for D/T will be derived as follows (Sahai 1993; Ulm 1990):

Lower confidence limit $L = \frac{0.5c_{\alpha/2,2D}}{T}$ for D>0, 0 otherwise,

Upper confidence limit $U = \frac{0.5c_{1-\alpha/2,2D+2}}{T}$

Where $c_{\alpha,k}$ is the αth quantile of the Chi-square distribution with k degrees of freedom.

5.6 Rule of exclusion criteria of analysis sets

Exclusion due to protocol deviation will be handled through the population attribute of the related estimand. The rules for participant classification in the analysis sets are based on criteria in Table 5-3.

Table 5-3 Criteria leading to exclusion

Analysis Set	Criteria that cause participants to be excluded
RAS	Not randomized
FAS	Not in RAN Mistakenly randomized and no double-blind study drug taken
SAF	No double-blind study drug taken
PK analysis set	Not in SAF No valid PK (LNA043) concentration measurement
ANGPTL3 analysis set	Not in RAN No valid ANGPTL3 concentration measurement
Immunogenicity Prevalence Set	Not in SAF No valid IG measurement at baseline or post-baseline

5.7 Table of Individual sub-regions for CCI analysis

Table 5-4 Cartilage Thickness Sub-regions

CCI

Table 5-5 Bone Marrow Lesion Sub-regions

CCI

6 Reference

Altman RD, Rosen JE, Bloch DA, et al (2009) A double-blind, randomized, saline-controlled study of the efficacy and safety of EUFLEXXA for treatment of painful osteoarthritis of the knee, with an open-label safety extension (the FLEXX trial). Semin Arthritis Rheum; 39(1):1-9

Bernard K, Grankov S, Van der Stoep M, et al (2020) FRI0393 Baseline characteristics of the study population in ROCCELLA, a phase 2 clinical trial evaluating the efficacy and the safety of S201086/GLPG1972 in patients with known osteoarthritis. Annals of the Rheumatic Diseases; 79:794-795

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

Collins NJ, Misra D, Felson DT, et al (2011) Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS). Arthritis Care Res (Hoboken); 63 Suppl 11(0 11):S208-28

Dougados M, Paternotte S, et al. (2010) ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/epidemiological studied in axial spondyloarthritis. Ann Rheum Dis; 70:249-251

Dworkin RH, Turk DC, Farrar JT, et al (2005) Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain; 113(1-2):9-19

FDA Guidance for Industry: Population Pharmacokinetics: https://www.fda.gov/media/71364/download.

Gandek B, Ware JE, Aaronson NK, et al (1998) Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol; 51(11):1171-8

Gandhi SK, Salmon JW, Zhao SZ, et al (2001) Psychometric evaluation of the 12-item short-form health survey (SF-12) in osteoarthritis and rheumatoid arthritis clinical trials. Clin Ther; 23(7):1080-98

Garwood F (1936). Fiducial limits for the Poisson distribution. Biometrika, 46; 441–453.Hochberg MC, Guermazi A, Guehring H, et al (2019) Effect of Intra-Articular Sprifermin

vs Placebo on Femorotibial Joint Cartilage Thickness in Patients With Osteoarthritis: The FORWARD Randomized Clinical Trial. JAMA; 322(14):1360-1370

ICH E9(R1) Harmonized Guideline: addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Final version on 20 November 2019.

McAlindon TE, Driban JB, Henrotin Y, et al (2015) OARSI Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for knee osteoarthritis. Osteoarthritis and Cartilage; 23(5):747-60

Manno RL, Bingham CO, Paternotte S, et al (2012) OARSI-OMERACT initiative: defining thresholds for symptomatic severity and structural changes in disease modifying osteoarthritis drug (DMOAD) clinical trials. Osteoarthritis and Cartilage; 20(2):93-101

Newcombe, RG. (1998) Two-sided confidence intervals for the single proportion: comparison of seven methods. Statistics in Medicine; 17: 857-872.

Permutt T (2016) Sensitivity analysis for missing data in regulatory submissions. Stat Med; 35(17):2876-9

Predictive margins and average marginal effects: http://support.sas.com/kb/63/038.html

Sahai H, Khurshid Anwer (1993). Confidence intervals for the mean of a poisson distribution: a review. Biom J, 35 (7); 857-867

Ulm K (1990). A simple method to calculate the confidence interval of a standard mortality ratio. American Journal of Epidemiology, 131(2); 373-375

Wolfe F, Clauw DJ, Fitzcharles MA, et al (2010) The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken); 62(5):600-10

Roger JH (2023) The RMConjPlus and RefBasedX SAS macro.

Carpenter JR, Roger JH and Kenward MG (2013) Analysis of Longitudinal Trials with Protocol Deviation: A Framework for Relevant, Accessible Assumptions, and Inference via Multiple Imputation. Journal of Biopharmaceutical Statistics, 23:6, 1352-1371.