STATISTICAL ANALYSIS PLAN

PROTOCOL NUMBER: ALXN1840-WD-205

A PHASE 2, SINGLE-ARM, PATHOLOGIST-BLINDED STUDY USING LIVER BIOPSY SPECIMENS TO ASSESS COPPER CONCENTRATION AND HISTOPA THOLOGIC CHANGES IN PATIENTS WITH WILSON DISEASE WHO ARE TREATED WITH ALXN1840 FOR 48 WEEKS FOLLOWED BY AN EXTENSION TREATMENT PERIOD WITH ALXN1840 FOR UP TO AN ADDITIONAL 48 WEEKS

Author: [redacted]

Date: 22 Feb 2023

Version: 2.0

1. APPROVAL SIGNATURES

[redacted]

21-Feb-2023 | 13:38 EST

[redacted] Date dd Mmm yyyy

[redacted]

21-Feb-2023 | 10:55 EST

[redacted] Date dd Mmm yyyy

[redacted]

21-Feb-2023 | 11:49 EST

[redacted] Date dd Mmm yyyy

[redacted]

21-Feb-2023 | 10:49 EST

[redacted] Date dd Mmm yyyy

2. TABLE OF CONTENTS, LIST OF TABLES, AND LIST OF FIGURES

TABLE OF CONTENTS

LIST OF TABLES

3. LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

The following abbreviations and acronyms are used in this Statistical Analysis Plan (SAP).

Table 1: Abbreviations and Acronyms

Abbreviation or Acronym	Explanation
ADL	activities of daily living
AE	adverse event
ALT	alanine aminotransferase
ALP	alkaline phosphatase
AST	aspartate transaminase
ATC	Anatomic Therapeutic Chemical
AUC	area under the plasma concentration versus time curve
AUC_t	area under the plasma concentration versus time curve from time 0 to the last quantifiable concentration
AUC_τ	area under the plasma concentration versus time curve over the dosing interval
AUEC	area under the effect versus time curve
$AUEC_{0\text{-}48W}$	daily mean area under the effect versus time curve of dNCC from 0 to 48 weeks
$AUEC_t$	area under the effect versus time curve from time 0 to the last quantifiable concentration time
BLQ	below the limit of quantification
BMI	body mass index
BPRS-24	Brief Psychiatric Rating Scale-24
CGI-I	Clinical Global Impression-Improvement Scale
CGI-S	Clinical Global Impression-Severity Scale
CI	confidence interval
CLDQ	Chronic Liver Disease Questionnaire
C_{max}	maximum observed concentration
cNCC	calculated non-ceruloplasmin-bound Cu
Cp	ceruloplasmin
CpC	ceruloplasmin-bound Cu
CRF	case report form
CS	clinically significant
CTCAE	Common Terminology Criteria for Adverse Events

Table 1: Abbreviations and Acronyms

Abbreviation or Acronym	Explanation
C_{trough}	trough (pre-dose) concentration observed at the start of the dosing interval
CV	coefficient of variation
dNCC	directly measured non-ceruloplasmin-bound Cu
ECG	electrocardiogram
eDISH	evaluation of drug-induced serious hepatotoxicity
E_{max}	maximum observed PD concentration in plasma
EoS	End of Study
EQ-5D	EuroQoL 5 Dimensions
EQ-5D-5L	EuroQoL 5 Dimensions 5 Levels
ET	Early Termination
FA	Full Analysis
FIB-4	Fibrosis-4
FSH	follicle-stimulating hormone
GGT	gamma-glutamyl transferase
GM	geometric mean
INR	international normalized ratio
LBC	labile-bound Cu
LLOQ	lower limit of quantification
ln	natural logarithm
LSM	least-squares mean
MedDRA	Medical Dictionary for Regulatory Activities
MELD	Model for End-Stage Liver Disease
MRE	magnetic resonance elastography
MR-PDFF	magnetic resonance-proton density fat fraction
NAFLD	nonalcoholic fatty liver disease
NAS	nonalcoholic fatty liver disease activity score
NCC	non-ceruloplasmin-bound Cu
NCI	National Cancer Institute
NCS	not clinically significant
PD	pharmacodynamic(s)

Table 1: Abbreviations and Acronyms

Abbreviation or Acronym	Explanation
PK	pharmacokinetic(s)
PP	Per Protocol
PT	Preferred Term
PTAE	pretreatment adverse event
PUF	plasma ultrafiltrate
QT	interval between the start of the Q wave and the end of the T wave in an ECG
QTcF	QT interval corrected using Fridericia's formula
REML	restricted maximum likelihood
SAE	serious adverse event
SAP	statistical analysis plan
SAS®	Statistical Analysis Software
SoA	schedule of activities
SOC	System Organ Class (MedDRA)
SoC	standard of care
SF-36	Short Form-36
TEAE	treatment-emergent adverse event
TPC	tripartite complex
TSQM-9	Treatment Satisfaction Questionnaire for Medication-9
ULOQ	upper limit of quantification
ULN	upper limit of normal
UWDRS	Unified Wilson Disease Rating Scale
VAS	visual analog scale
WBC	white blood cell
WD	Wilson disease
WHO	World Health Organization

4. DESCRIPTION OF THE PROTOCOL

This is a Phase 2, single-arm, pathologist-blinded study, using liver biopsy specimens to assess copper (Cu) concentration and histopathologic changes in patients with Wilson disease (WD), who are treated with ALXN1840 for 48 weeks followed by an Extension Treatment Period with ALXN1840 for up to an additional 48 weeks. This study is being conducted to evaluate the change from Baseline in liver Cu concentration at 48 weeks in adult patients with WD treated with ALXN1840 at doses up to 30 mg who have been treated for at least 1 year with standard of care (SoC; trientine, penicillamine, and zinc). Currently available SoC medications (e.g., trientine, penicillamine, and zinc) control excess free Cu in the blood but have never been demonstrated prospectively to reduce Cu burden in the human liver.

Approximately 28 patients with WD who are aged 18 years and older will be enrolled into the study globally, such that approximately 24 evaluable patients complete the 48-week Treatment Period.

The planned study duration is approximately 100 weeks, including an up to 4-week Screening Period. A liver biopsy will be performed during the Screening Period up to 4 weeks before the first dose of study drug, followed by a 48-week Treatment Period, which starts on the day of the first dose of study drug. Patients who complete the 48-week Treatment Period will be offered the opportunity to participate in a 48-week Extension Period to evaluate the long-term safety and efficacy of ALXN1840. For patients who enter the Extension Period, Week 48 Visit is the end of the Treatment Period and the beginning of the Extension Period (i.e., the Week 48 Visit and the Extension Period Day 1 visit will occur on the same day). For patients who enter the Extension Period, all assessments for the Week 48 Visit and the liver biopsy must be performed before the first dose of ALXN1840 on Day 1 of the Extension Period. Dosing of ALXN1840 on Extension Period Day 1 marks the beginning of the Extension Period. Patients who do not enter the Extension Period will discontinue dosing at Week 48 following the second liver biopsy and will have a final study visit for safety follow up at Week 52 of the Treatment Period. The schedule of activities is available in Section 9.1. All biopsy specimens will be reviewed and scored by an independent expert hepatopathologist who is blinded to patient identity and biopsy order. The protocol can be referenced for additional details.

The analysis and statistical reporting will be conducted using SAS® version 9.4 or higher (SAS Institute Inc., Cary, North Carolina).

4.1. Changes From Analyses Specified in the Protocol

The following changes were made from analyses specified in the Protocol Amendment (approval dated 27 May 2022):

For exploratory endpoints measured only at Baseline and Week 48, a paired sample t-test or a Wilcoxon Signed-rank test will be used to assess Baseline and Week 48 results for continuous or ordinal outcomes, respectively, in this SAP per Section 7.2.4.

These analyses are to replace the ANCOVA or nonparametric ANCOVA models mentioned in Protocol Amendment 5 (approval date 27 May 2023) Section 9.4.1.1 for these endpoints.

4.2. Changes From Analyses Specified in the Previous Version of the Statistical Analysis Plan

The following updates have been made to the SAP version 1.0 approved on 16 Dec 2022:

Section No. and Name	Description of Change	Justifications
Section 3. Table 1. List of Abbreviations and Definitions of Terms	Added the following abbreviations: • cNCC for calculated non-ceruloplasmin-bound Cu • dNCC for directly measured non-ceruloplasmin-bound Cu • $AUEC_{0\text{-}48W}$ for daily mean AUEC of dNCC from 0 to 48 weeks • FIB-4 for Fibrosis-4 • PT for Preferred Term • ULOQ for upper limit of quantification	To align with Protocol Amendment 5
Section 4. Description of the Protocol	Clarified that a liver biopsy is needed for patients entering Extension Period on Week 48 Clarified that all biopsy samples will be reviewed and scored by an independent expert hepatopathologist blinded to identity and biopsy order	To align with Protocol Amendment 5
Section 4.1. Changes From Analyses Specified in the Protocol	Added statements that ANCOVA or nonparametric ANCOVA models are replaced with a paired sample t-test or a Wilcoxon signed-rank test for continuous exploratory endpoints or ordinal exploratory endpoints tested only at Baseline and Week 48	To align with executed analysis and learnings for this study
Section 4.2. Changes From Analyses Specified in the Previous Version of the Statistical Analysis Plan	Provided summary of changes made in this SAP relative to SAP version 1.0 (approval date 16 Dec 2022)	Good documentation practice
Section 5.1.3. The Exploratory Efficacy Endpoints	Added the following exploratory endpoints: • Absolute and percent change in directly measured non-ceruloplasmin-bound Cu (dNCC) and labile-bound Cu (LBC) from Baseline to Week 48 • Daily mean AUEC of dNCC from 0 to 48 weeks	To align with Protocol Amendment 5

Section No. and Name	Description of Change	Justifications
	Corrected NCC to cNCC Aligned text of MR-PDFF with protocol by replacing steatosis with liver stiffness	
Section 5.1.3.10. Fibrosis-4 Index	Newly added section to define Fibrosis-4 (FIB-4)	To align with executed analysis and learnings for this study
Section 5.2.1. Adverse Events	Updated protocol approval date from 20 Dec 2019 to 27 May 2022	To align with Protocol Amendment 5
Section 5.3. Pharmacokinetics/Pharmacodynamics	Added PD endpoints of plasma concentration of directly measured non-ceruloplasmin-bound Cu (dNCC), LBC, and AUEC of dNCC Corrected NCC to cNCC	To align with Protocol Amendment 5
Section 5.4. Biomarkers	Clarified that it is 24-hour urine samples to be collected per the SoA	Administrative updates
Section 6.2. Data Sets in the Treatment Period Section 6.3. Data Sets in the Extension Period	Added “evaluable” after “liver Cu value”	Clarification
Section 6.2. Data Sets in the Treatment Period Section 6.3. Data Sets in the Extension Period	Added “evaluable” after “liver Cu value”	Clarification
Section 7.1.3. Demographics, Disease Characteristics, and History	Removed statement of “Demographic and Baseline disease characteristics will also be produced for the subgroups mentioned in Section 7.2.1.2”	Subgroup analysis is considered unnecessary.
Section 7.1.3.1. Demographics	Added the following demographic variables: • Time since WD diagnosis (months) • Cumulative duration of prior WD treatment (months) • Prior WD therapy • Cirrhosis	To align with executed analysis and learnings for this study

Section No. and Name	Description of Change	Justifications
Section 7.1.3.2. Disease Characteristics	Added the following disease characteristics variables: • Liver ultrasound • dNCC (µmol/L) • Creatinine (µmol/L) • High-density lipids (mmol/L) • Low-density lipids (mmol/L) • Triglycerides (mmol/L) • 24-hour urinary creatinine (µmol/day) Corrected the following disease characteristics variables: • NCC to cNCC • 24-hour urinary Cu and Mo unit to "ug/day" from "ug/mL"	To align with Protocol Amendment 5 To align with executed analysis and learnings for this study
Section 7.2.1.1. Handling of Dropouts or Missing Data	Added following note: Liver Cu concentrations reported as < LLOQ or > ULOQ will be considered non-evaluable. Non-evaluable data will be excluded when deriving summary statistics in both liver Cu primary analysis and Cu sensitivity analysis. This same principle will also be applied to the analysis of other live metals	Clarifications
Section 7.2.1.2. Subgroup Analyses	Updated subgroup analyses are considered unnecessary	To align with Protocol Amendment 5
Section 7.2.1.5. Sensitivity Analyses	Clarified to use observation data from patients where both Baseline and Week 48 data are available	To add specifics for sensitivity analysis using observed data
Section 7.2.4. Other Efficacy Analyses	Corrected NCC to cNCC Added dNCC and LBC Added liver metals from liver biopsy specimen for a paired sample t-test at Baseline and Week 48 when applicable	To align with Protocol Amendment 5

Section No. and Name	Description of Change	Justifications
Section 7.2.4. Other Efficacy Analyses	Added FIB-4	To align with executed analysis and learnings for this study
Section 7.2.5.1. Pharmacokinetic and Pharmacodynamic Sampling	Corrected NCC to cNCC Added dNCC and LBC	To align with Protocol Amendment 5
Section 7.2.8. Pharmacodynamic and Biomarker Concentration Analyses	Corrected NCC to cNCC Added dNCC Defined dNCC, AUEC of dNCC, daily mean AUEC of dNCC from 0 to Week 48 ($AUEC_{0\text{-}48W}$), and corresponding analysis plans	To align with Protocol Amendment 5
Section 7.3.3.2 eDISH	Removed “each treatment arm (1840 and SoC) separately” Corrected “eDISH Plot by Treatment Arm” to “eDISH plot for all patients"	Correction. All patients received 1840
Section 7.5. Interim Analyses	Clarified that interim analyses may include safety data only or safety and efficacy data together Clarified that interim analysis outputs may include a full or a subset of listed outputs when appropriate	To align with executed analysis and learnings for this study
Section 9.1 Protocol Schedule of Events	Removed schedule of assessment table and provide reference to Protocol Amendment 5 Section 1.3 Table 1 and Table 2	To align with Protocol Amendment 5
Section 9.4.12 CLDQ	Added specifics regarding CLDQ scoring principles	To clarify how CLDQ scores are derived
All Sections	Replaced copper with Cu and molybdenum with Mo throughout the text when applicable Formatting updates per style guide Spelling and grammar updates	Administrative updates

5. DEFINITIONS

5.1. Efficacy

5.1.1. Primary Endpoint(s)

The primary efficacy endpoint is the change from Baseline to Week 48 in liver Cu concentration.

5.1.2. Secondary Endpoints

The secondary efficacy endpoints are as follows:

- Change from Baseline to Week 48 in steatosis, inflammation, and fibrosis
- Change from Baseline in Clinical Global Impression-Severity Scale (CGI-S)
- Change from Baseline in Clinical Global Impression-Improvement Scale (CGI-I)

5.1.2.1. Steatosis, Inflammation, and Fibrosis

Assessments will be based, in part, on a histological scoring system for nonalcoholic fatty liver disease (NAFLD).

Inflammation will be quantified by the NAFLD activity score (NAS). The score is defined as the unweighted sum of the scores for steatosis (0 to 3), lobular inflammation (0 to 3), and hepatocellular ballooning (0 to 2), thus ranging from 0 to 8 (Kleiner, 2005).

Steatosis from histology will be evaluated in 2 ways: the steatosis component of the NAS, expressed as an ordinal score ranging from 0 to 3, and morphometric quantification of hepatic fat, expressed as a percentage (Jayakumar, 2018).

Fibrosis from histology will be evaluated in 2 ways: graded on an ordinal scale from 0 to 4 (METAVIR, 1994) and by morphometric quantification of hepatic collagen content and alpha-smooth muscle actin (a-SMA), expressed as percentages.

5.1.2.2. CGI-I and CGI-S

The CGI-S is a 7-point scale clinician assessment. Patients are assessed on severity of illness at the time of rating/assessment as follows: 1, normal, not at all ill; 2, borderline ill; 3, mildly ill; 4, moderately ill; 5, markedly ill; 6, severely ill; or 7, extremely ill.

The CGI-I is a 7-point scale clinician assessment. Patients are assessed as improved or worsened relative to a Baseline state at the beginning of the intervention and rated as follows: 1, very much improved; 2, much improved; 3, minimally improved; 4, no change; 5, minimally worse; 6, much worse; or 7, very much worse.

5.1.3. The Exploratory Efficacy Endpoints

The exploratory efficacy endpoints are the following:

- Change from Baseline to Week 48 in Model for End-Stage Liver Disease (MELD) score and modified Nazer score
- Liver stiffness by magnetic resonance elastography (MRE)

- Liver stiffness by transient elastography (vibration-controlled transient elastography/FibroScan)
- Liver stiffness by magnetic resonance-proton density fat fraction (MR-PDFF)
- Mitochondrial ultrastructure by electron microscopy
- Cu deposition in the eye by optical coherence tomography
- Absolute and percent change in plasma total and plasma ultrafiltrate (PUF) Cu and calculated non-ceruloplasmin-bound Cu (cNCC) concentrations, as well as directly measured NCC (dNCC), LBC, ceruloplasmin (Cp), and Cp-bound Cu (CpC) from Baseline to Week 48
- Daily mean area under the effect versus time curve (AUEC) of dNCC from 0 to 48 weeks
- Change from Baseline to Week 48 in quality-of-life measures: EuroQoL 5 Dimensions 5 Levels (EQ-5D-5L), Short Form-36 (SF-36), and Chronic Liver Disease Questionnaire (CLDQ)
- Treatment Satisfaction Questionnaire for Medication-9 (TSQM-9)
- Change from Baseline to Week 48 in Brief Psychiatric Rating Scale-24 (BPRS-24) and Unified Wilson Disease Rating Scale (UWDRS)
- Change from Baseline in the Fibrosis-4 (FIB-4) index

5.1.3.1. MELD Score

The MELD is a scoring system for assessing the severity of chronic liver disease in adults and adolescents 12 years and older. The MELD score ranges from 6 to 40, with higher values indicating more advanced disease. Refer to Appendix 9.4.10 for details on calculating the MELD score.

5.1.3.2. Modified Nazer Score

The modified Nazer score is an assessment of liver status and consists of a composite of 5 laboratory parameters: AST, international normalized ratio (INR), bilirubin, albumin, and white blood cell (WBC) count. The score has a total range of 0 to 20, and lower values indicate a healthier liver status. Refer to Appendix 9.4.11 for details on calculating the modified Nazer score.

5.1.3.3. Imaging Techniques

Imaging techniques, including MRE, transient elastography, MR-PDFF, electron microscopy, and optical coherence tomography, will be detailed in an imaging manual. Steatosis by imaging will be measured by MR-PDFF and expressed as a percentage. Liver stiffness will be assessed by both MRE and transient elastography (FibroScan) measured in kPa (Jayakumar, 2018).

5.1.3.4. EQ-5D

The EuroQoL 5 Dimensions (EQ-5D) consists of 2 different assessments: the EQ-5D-5L descriptive system and the EQ-5D visual analog scale (VAS). The descriptive system comprises

measures of health-related quality-of-life state and consists of 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 levels of severity: no problems, slight problems, moderate problems, severe problems, or extreme problems. The EQ VAS is a simple score the patient is asked to complete to record the patient's self-rated health on a vertical VAS. The EQ-5D response profiles will be converted into a single number called index score, using the UK 3L crosswalk value set.

Both the VAS score and the index score will be endpoints for analysis. The EQ-5D will be administered at visits listed in Section 9.1.

5.1.3.5. SF-36

The SF-36 is a 36-item survey that is used to assess the patient's view of their health. It consists of 8 scaled scores that are weighted sums of the questions in each section (Vitality, Physical functioning, Bodily pain, General health perceptions, Physical role functioning, Emotional role functioning, Social role functioning, Mental health). The lower the score, the greater the disability. The scores from the 8 subscales measurements will be converted by QualityMetric® Scoring System to 2 summary scores: the physical component summary and the mental component summary. There is no single overall score for the SF-36.

5.1.3.6. CLDQ

The CLDQ is a 29-item instrument with 6 domains (Fatigue, Activity, Emotional function, Abdominal symptoms, Systemic symptoms, Worry) that measures longitudinal change over time in patients with chronic liver disease. Each domain is scored from 1 (worst possible function) to 7 (best possible function). Refer to Appendix 9.4.12 for details on calculating CLDQ.

5.1.3.7. TSQM-9

The abbreviated TSQM-9 is used to assess the overall level of satisfaction or dissatisfaction with medication patients are taking. Refer to Appendix 9.4.13 for details on calculating TSQM-9.

The TSQM-9 will be administered at visits listed in Section 9.1.

5.1.3.8. BPRS-24

The BPRS-24 is a 24-item instrument that allows the rater to measure psychopathology severity. The presence and severity of psychiatric symptoms are rated on a Likert scale ranging from 1 (not present) to 7 (extremely severe).

The BPRS-24 will be obtained at visits listed in Section 9.1. The total score across the 24 items will be endpoint for analysis. A minimum of 20 out of 24 items are required to be completed at any given visit; the total score will then be taken for the non-missing items and scaled up to the 24-item score for the purposes of data summary and analysis. If fewer than 20 of 24 items are completed at any given visit, the BPRS-24 total score will be set to missing for that patient visit.

5.1.3.9. Unified Wilson Disease Rating Scale (Parts I, II, and III)

The UWDRS is a clinical rating scale designed to evaluate the neurological manifestations of WD. The UWDRS comprises 3 parts: UWDRS Part I (consciousness, item 1), UWDRS Part II (a

historical review of daily activity items, items 2 to 11 [10 items in total]), and UWDRS Part III (a neurological examination, items 12 to 34 [23 items in total]).

The UWDRS will be assessed at visits listed in Section 9.1. The total score for Part I, Part II, Part III, the following subscales and individual symptom scales: Part II Activities of Daily Living (ADL), Part II Salivation, Part II Swallowing, Part III Tremor subscale, Part III Dystonia subscale, Part III Speech, and Part III Handwriting will be summarized.

Refer to Appendix 9.4.14 for details on calculating UWDRS.

5.1.3.10. Fibrosis-4 Index

The FIB-4 Index (Vallet-Pichard, 2007) is a formula used to predict liver fibrosis based on standard biochemical values (alanine aminotransferase [ALT], aspartate aminotransferase [AST], and platelet count) and age. The FIB-4 Index will be calculated by a Central Laboratory. The formula is as follows:

$$[\text{Age (years)} \times \text{AST (U/L)}] / [\text{Platelets } (10^9/\text{L}) \times \text{sqrt}(\text{ALT } \{\text{U/L}\})]$$

The FIB-4 Index will be assessed at Day 1, Week 4, Week 6, Week 8, Week 12, Week 18, Week 24, Week 36, Week 48/ET, and EOS. If any of the parameters required to compute the FIB-4 Index are unavailable at any given visit, FIB-4 Index will be set to missing for that patient visit.

5.2. Safety

The safety endpoints in this study are the safety and tolerability of individualized dosing of ALXN1840 over time. Long-term safety and tolerability of ALXN1840 will be obtained in the Extension Period. Safety will, in part, be assessed by analysis of adverse events (AEs). The AE data will be collected regarding onset, duration, seriousness, intensity, outcome and relatedness to study drug. Deaths and other serious adverse events (SAEs) will also be evaluated and will be collected on a separate case report form (CRF). Changes in physical examinations, vital signs (resting heart rate, semi-supine systolic and diastolic blood pressure, respiratory rate, temperature, and weight), and laboratory values will also be evaluated and assessed.

Planned time points for all safety endpoints are provided in the schedule of activities (SoAs) in Section 9.1.

5.2.1. Adverse Events

AEs are defined in Section 10.3 of the Protocol (dated 27 May 2022). Details on how to define treatment-emergent adverse events (TEAEs) with partial dates and on how to handle missing or partial AE dates are described in Section 9.4.1.

The occurrence of AEs at each visit will be recorded on designated CRF pages. Each AE is to be characterized (i.e., verbatim term) and information will be provided regarding its seriousness, start and stop dates, intensity, outcome, and causal relationship with the study drug. The safety evaluation will include an assessment of all AEs, SAEs, AE intensity, and AE causality. AE intensity will be evaluated using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 5.0 (published 27 Nov 2017). AE causality will be

evaluated by Investigators to be either not related or related. AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 23.0 or higher and summarized by primary System Organ Class (SOC) and Preferred Term (PT).

5.2.2. Vital Signs

Vital signs will be measured in a semi-supine position after 5-minute rest and will include temperature, systolic and diastolic blood pressure, pulse, and respiratory rate. Blood pressure and pulse measurements should be preceded by at least 5 minutes of rest for the patient in a quiet setting without distractions (e.g., television and cell phones). Weight and height will also be measured.

5.2.3. Electrocardiogram

Single 12-lead Electrocardiograms (ECGs) will use an ECG machine that automatically calculates the heart rate and measures PR, QRS, RR, interval between the start of the Q wave and the end of the T wave in an ECG (QT), and QT interval corrected using Fridericia's formula (QTcF) intervals. Patients must be supine for approximately 5 to 10 minutes before ECG collection and remain supine but awake during ECG collection.

5.2.4. Clinical Laboratory Assessments

Clinical chemistry, hematology, coagulation, urinalysis, and other tests will be performed. For females of childbearing potential, a serum or urine pregnancy test will be performed. For post-menopausal females, follicle-stimulating hormone (FSH) will be performed. The list of specific tests is provided in Section 10.2 of the Protocol.

5.2.5. Physical Examinations

A complete physical examination will include, at a minimum, assessments of cardiovascular, respiratory, gastrointestinal, and neurological systems. A brief physical examination will include, at a minimum, assessments of the skin, lungs, cardiovascular system, and abdomen (liver and spleen).

5.2.6. Other Safety Events of Special Interest

No other safety events of special interest have been defined for this study.

5.3. Pharmacokinetics/Pharmacodynamics

At study visits indicated in the SoAs (Section 9.1), the following samples will be collected for all patients:

- Blood samples will be collected for measurement of total and PUF Mo in plasma.
- Blood samples will be collected for measurement of total and PUF Cu, cNCC, dNCC, and LBC in plasma. Daily mean AUEC of dNCC from 0 to 48 weeks will also be derived.
- A part of the liver tissue sample collected for biopsy will be allocated for the measurement of total Mo, Cu, and other metals concentrations in the liver tissue when available.

5.4. Biomarkers

Collection of samples for other biomarker research is also part of this study. At study visits indicated in the SoAs (Section 9.1), the following samples will be collected for all patients:

- Blood samples will be collected to measure Cp and CpC.
- 24-hour urine samples will be collected for the analysis of urine creatinine, Cu, and Mo.

6. DATA SETS ANALYZED (STUDY POPULATIONS)

The Treatment Period and Extension Period will be analyzed separately and sequentially.

6.1. Enrolled

All patients who sign the informed consent form and passed all inclusion/exclusion criteria will be considered as enrolled.

6.2. Data Sets in the Treatment Period

6.2.1. Safety Analysis Set in the Treatment Period

Safety Analysis Set in the Treatment Period is defined as all patients who receive at least 1 dose of treatment in the Treatment Period. This population will be used for the safety analyses in the Treatment Period.

6.2.2. Full Analysis (FA) Set in the Treatment Period

The Full Analysis (FA) Set in the Treatment Period is defined as all patients who receive at least 1 dose of study drug in the Treatment Period and have at least a Baseline liver Cu value evaluable. This population will be used for the primary, secondary, and exploratory efficacy analyses.

6.2.3. Per Protocol (PP) Set in the Treatment Period

The Per Protocol (PP) Set in the Treatment Period is defined as all patients who receive at least 1 dose of study drug in the Treatment Period and have both Baseline and Week 48 liver Cu values evaluable. Patients with important protocol deviations that are likely to impact the primary efficacy endpoint analysis will be excluded from the PP Set. Important protocol deviations and the PP Set will be defined, documented, and agreed within Alexion prior to database lock. This population will be used for the sensitivity analyses for the primary endpoint.

6.2.4. Pharmacokinetic (PK) and Pharmacodynamic (PD) Sets in the Treatment Period

Patients meeting the definition for the Safety Analysis Set in the Treatment Period (Section 6.2.1) and having evaluable PK data for total Mo and/or PUF Mo (as surrogate measure for ALXN1840 PK) in plasma in the Treatment Period will be included in the PK Set in the Treatment Period.

Patients meeting the definition of Safety Analysis Set in the Treatment Period (Section 6.2.1) and having evaluable PD data for total Cu and/or PUF Cu in plasma in the Treatment Period will be included in the PD Set in the Treatment Period.

6.3. Data Sets in the Extension Period

6.3.1. Safety Analysis Set in the Extension Period

Safety Analysis Set in the Extension Period is defined as all patients who receive at least 1 dose of ALXN1840 in the Extension Period. This population will be used for the safety analyses in the Extension Period.

6.3.2. Full Analysis (FA) Set in the Extension Period

The FA Set in the Extension Period is defined as all patients who receive at least 1 dose of ALXN1840 in the Extension Period and have at least a Baseline liver Cu value evaluable. This population will be used for the efficacy analyses in the Extension Period.

6.3.3. Pharmacokinetic (PK) and Pharmacodynamic (PD) Sets in the Extension Period

Patients meeting the definition of Safety Analysis Set in the Extension Period (Section 6.3.1) and having evaluable PK data for total Mo and/or PUF Mo (as surrogate measure for ALXN1840 PK) in plasma in the Extension Period will be included in the PK Set in the Extension Period.

Patients meeting the definition of Safety Analysis Set in the Extension Period (Section 6.3.1) and having evaluable PD data for total Cu and/or PUF Cu in plasma in the Extension Period will be included in the PD Set in the Extension Period.

7. STATISTICAL ANALYSIS

The Treatment Period and Extension Period will be analyzed separately and sequentially. When summarizing by descriptive statistics, continuous variables will be summarized using the number of non-missing values, mean, standard deviation (SD), median, minimum, and maximum. Certain continuous variables may also include 95% confidence intervals (CIs), and certain categorical variables will be summarized as proportions with 95% CIs. Categorical variables will be summarized using percentages and frequency counts. Patient level listings will be provided accordingly without separating the Treatment Period and Extension Period.

Baseline for all assessments will be defined as the last assessment yielding non-missing valid values taken prior to the first dose of ALXN1840. Details can be found in Appendix 9.4.

The analysis and statistical reporting will be conducted using SAS® version 9.4 or higher (SAS Institute Inc., Cary, North Carolina).

7.1. Study Patients

7.1.1. Disposition of Patients

An overview of patient populations will be summarized, where the number and percentage of patients in each analysis set will be presented.

Frequency counts and percentages of screen failure reasons will be provided for patients who failed to meet study entry requirements and patients who failed not due to eligibility criteria during Screening.

The number and percentage of patients who completed, or prematurely discontinued from the study will be described. For patients who discontinued the study, the number and percentage will be summarized by their reasons for premature discontinuation. A summary will be provided for patients by region and country.

Additionally, a summary of patients who did not meet inclusion or who met exclusion criteria will be provided.

Descriptive statistics of the number of days in study will be summarized. The date of the first and last use of study drug in each period and the study termination date will be listed.

A listing of patients will be provided for all enrolled patients, with the extent of their participation in the study and the reason for discontinuation. A listing of analysis populations for all enrolled patients and a listing of screen failure patients will also be provided. Additionally, a listing of patients and the inclusion criteria they failed to meet and the exclusion criteria they met will be provided.

7.1.2. Protocol Deviations

All important/not important protocol deviation will be determined and appropriately categorized prior to database lock.

The number and percentage of patients with any important/not important protocol deviations as well as the number and percentage of patients with deviations within each category will be presented. Corresponding listing will be provided to include the category and important/not

important status. A listing of the reasons for which patients were excluded from efficacy analyses, including the PP Set, will also be provided.

7.1.3. Demographics, Disease Characteristics, and History

All demographic, Baseline disease characteristics, medical/surgical history, and Baseline physical examination will be summarized for the FA Set in the Treatment Period. Continuous variables will be presented using descriptive statistics, and categorical variables will be presented using frequencies and percentages. Age will be calculated relative to date of informed consent and will be summarized as both continuous and categorical variables. Listings will also be provided.

7.1.3.1. Demographics

The following demographic variables will be summarized:

- Sex
- Race
- Ethnicity
- Country
- Region
- Age (years) at informed consent
- Age categories at informed consent (18 to <65 years and ≥ 65 years)
- Height (cm)
- Weight (kg)
- Body mass index (BMI) (kg/m^2)
- Time since WD diagnosis (months)
- Cumulative duration of prior WD treatment (months)
- Prior WD therapy
- Cirrhosis

The BMI will be calculated based on collected height and weight. Details are in Section 9.4.7.

7.1.3.2. Disease Characteristics

The following Baseline disease characteristics will be summarized:

- Live ultrasound
- LBC (µmol/L)
- LBC categorization based on normal reference range (female: < 0.7 µmol/L, 0.7 to 5.9 µmol/L, and > 5.9 µmol/L; male: < 0.9 µmol/L, 0.9 to 4.4 µmol/L, and > 4.4 µmol/L)
- cNCC (µmol/L)

- dNCC (µmol/L)
- cNCC categorization based on normal reference range (< 0.8 µmol/L, 0.8 to 2.3 µmol/L, and > 2.3 µmol/L)
- UWDRS Part II
- UWDRS Part III
- CGI-S score
- MELD score
- Modified Nazer score
- EQ-5D score and EQ-5D VAS
- SF-36 score
- CLDQ score
- TSQM-9 score
- BPRS-24 score
- Albumin (g/L)
- Total bilirubin (µmol/L)
- Direct bilirubin (µmol/L)
- Alanine aminotransferase (ALT) (U/L)
- Aspartate transaminase (AST) (U/L)
- Gamma glutamyl transferase (GGT) (U/L)
- Platelets (10^9/L)
- WBC (also known as leukocytes) count (10^9/L)
- Neutrophils (10^9/L)
- Creatinine (µmol/L)
- Creatine phosphokinase (U/L)
- High-density lipids (mmol/L)
- Low-density lipids (mmol/L)
- Triglycerides (mmol/L)
- Total Cu (ng/mL)
- Total Mo (ng/mL)
- PUF Cu (ng/mL)
- PUF Mo (ng/mL)

- Ceruloplasmin (Cp)(mg/L)
- 24-hour urinary Cu (µg/day)
- 24-hour urinary Mo (µg/day)
- 24-hour urinary creatinine (µmol/day)

7.1.3.3. Medical/Surgical History and Baseline Physical Examination

Medical and surgical history will be summarized by counts and percentages and displayed by SOC and PT within each SOC. SOC and PT will be coded using the MedDRA version 23.0 or higher. This dictionary will be used throughout the life of the study and will not be updated during study conduct. A by-patient listing with the medical history verbatim text, symptom start date, end date, and ongoing/previous conditions will be presented.

Time since WD treatment start date (months) will be calculated as months occurring between enrollment and the start date from WD treatment history CRF and will be summarized using descriptive statistics and all details of WDs diagnosis will be listed in full.

For the physical examination, the number and percentage of patients with abnormal findings at Baseline for each body system will be summarized and included in a by-patient listing.

7.1.4. Prior and Concomitant Medications/Therapies

The World Health Organization (WHO) Drug Global version from Mar 2020 or later will be used to code the prior or concomitant medications/therapies. Medications will be summarized by Anatomic Therapeutic Chemical (ATC) Level 3 class and generic drug name.

Prior and concomitant medications will be summarized separately for the FA Set in the Treatment Period. The number and percentage of patients receiving any concomitant medication will be summarized, as well as the number and percentage of patients receiving any concomitant medication by ATC drug class and generic drug name. Patients reporting use of > 1 medication at each level of summarization (any medication received, ATC class, and generic drug name) will be counted only once. The ATC class terms will be displayed by descending order of incidence, as well as generic drug names within each ATC class. Prior medications used to treat WD and an additional analysis of all prior medications will be summarized similarly. Prior medications will be defined as medications that were discontinued prior to the start of study drug. Concomitant medications will be defined as medications that either started prior to the first dose of study drug and were continuing at the time of the first dose of study drug or started on or after the date of the first dose of study drug. If it cannot be determined whether a medication was stopped prior to the start of study drug dosing due to partial or missing medication start or end dates, it will be considered a concomitant medication. Further details on partial or missing dates can be found in Appendix 9.4.

Prior and concomitant medications will be presented in a listing by patient and medication name.

7.2. Efficacy Analyses

All efficacy analyses and summaries will be based on the FA Set in the Treatment Period. Additionally, the primary analysis will be performed on the PP Set in the Treatment Period. A 2-sided Type I error rate of 0.05 will be used to define statistical significance for all endpoints

without adjusting for multiplicity. Sample SAS code for the following analyses can be found in Section 9.5.

7.2.1. Primary Analysis

The primary objective of this study is to evaluate change from Baseline in liver Cu concentration following treatment with ALXN1840 at Week 48 in patients with WD. The primary analysis will be performed using the FA Set in the Treatment Period.

To assess the reduction in liver Cu concentration, liver biopsies will be taken during the Screening Period and at Week 48 of Treatment Period, where the assessment from Screening will be regarded as Baseline. The change from Baseline in liver Cu concentration will be evaluated using a paired sample t-test on Baseline and Week 48 assessment values.

Prior to the primary analysis, multiple imputation will be used to impute missing data at Week 48 due to any reason based on Baseline values. Missing values at Week 48 will be imputed 20 times via SAS PROC MI, using the MCMC method, including Baseline and Week 48 liver Cu assessment values. A paired sample t-test analysis will be applied for each of the 20 imputed datasets, and then 20 sets of analysis results will be combined afterward via SAS PROC MIANALYZE. The combined estimate of the treatment effect, standard error (SE), 95% CI, and p-value will be summarized.

7.2.1.1. Handling of Dropouts or Missing Data

For missing liver Cu concentration data at Week 48, the multiple imputation method will be used for the primary analysis and sensitivity analysis. Details for multiple imputation can be referred to in Section 7.2.1, and an example of SAS code can be referred to in Section 9.5. The Baseline value carried forward method will also be used in the second sensitivity analyses, where missing Week 48 values will be imputed to Baseline values.

Note: Liver Cu concentrations reported as < LLOQ or > ULOQ will be considered non-evaluable. Non-evaluable data will be excluded when deriving summary statistics in both liver Cu primary analysis and Cu sensitivity analysis. This same principle will also be applied to the analysis of other live metals.

7.2.1.2. Subgroup Analyses

Subgroup analyses are considered unnecessary. Additional subgroup analyses may be performed post-hoc, as appropriate.

7.2.1.3. Multicenter Studies

This is a multicenter study. Efficacy data collected from all study centers will be pooled for data analysis.

7.2.1.4. Hypothesis Testing and Significance Level

The hypothesis to be tested for the primary endpoint is as follows:

$$H_0: \mu_{ALXN1840} \geq 0 \text{ versus } H_1: \mu_{ALXN1840} < 0$$

where $\mu_{ALXN1840}$ is the mean change from Baseline in liver Cu concentration (µg Cu/g dry liver) at Week 48. A 2-sided Type I error rate of 0.05 will be used to define statistical significance without adjusting for multiplicity.

7.2.1.5. Sensitivity Analyses

Three sensitivity analyses will be performed to assess the robustness of the analysis result, as follows:

- Similar to the primary analysis but use Baseline value carried forward to impute missing Week 48 data instead
- Similar to the primary analysis but use the observed data from patients where both Baseline and Week 48 data are available instead
- Similar to the primary analysis but use the PP Set in the Treatment Period instead

7.2.2. Secondary Analyses

All secondary analyses will be performed using the FA Set in the Treatment Period.

Histologic assessments (steatosis, inflammation, and fibrosis) and CGI-S will be summarized for Baseline Visit and post-baseline visits using descriptive statistics. Change from Baseline will also be summarized for each post-baseline visits. Since CGI-I is only collected post-dose, only post-baseline visits will be summarized. In addition, the following analysis will be done.

The CGI-S data will be analyzed using MMRM. In the MMRM model, the response variables will be change from Baseline. Visit and Baseline values will be used as covariates. The restricted maximum likelihood (REML) estimation will be used. An unstructured covariance matrix will be used to model the within-subject error and the Kenward-Roger approximation will be used to estimate the degrees of freedom. If the fit of the unstructured covariance structure fails to converge or a positive definite Hessian matrix is not produced, the following covariance structures will be tried in order until convergence is reached: Toeplitz with heterogeneity, autoregressive with heterogeneity, Toeplitz, autoregressive, and compound symmetry. Based on the model, the comparison of post-baseline versus. Baseline time points will be tested using a 2-sided test at the alpha = 0.05 level of significance. The least squares mean (LSM) change from Baseline and associated SE will also be presented along with the 95% CI. The reference SAS code for the MMRM model can be found in Section 9.5.

Mean and 95% CI of CGI-S will be plotted over time. Additionally, patient level data over time will be plotted by spaghetti plot.

7.2.3. Tertiary Analyses

Not applicable.

7.2.4. Other Efficacy Analyses

All exploratory analyses will be performed using the FA Set in the Treatment Period.

Assessment for each exploratory endpoint will be summarized for Baseline and post-baseline visits by descriptive statistics. Change from Baseline and percent change from Baseline (for plasma total and PUF Cu, cNCC, dNCC, LBC, Cp, and CpC only) will also be summarized for each post-baseline visits. In addition, following analyses will be done.

Endpoints with continuous outcome, which are measured multiple times between Baseline and Week 48, will be analyzed using the MMRM in the same manner as of CGI-S. Details of the MMRM can be referred to in Section 7.2.2.

- MELD
- Modified Nazer
- Plasma total and PUF Cu
- cNCC
- dNCC
- LBC
- Cp and CpC
- EQ-5D
- SF-36
- CLDQ
- BPRS-24
- FIB-4
- 24-hour Urine Cu/Mo/creatinine

For plasma total and PUF Cu, cNCC, dNCC, LBC, Cp, and CpC, additional MMRM model will be performed on percent change from Baseline.

For all exploratory endpoints with multiple measurements between Baseline and Week 48, the following plots will also be provided: mean and 95% CI plot over time, and spaghetti plot for patient level data over time.

For endpoints with continuous outcome that are only measured at Baseline and Week 48, a paired sample t-test on Baseline and Week 48 assessment values will be performed when data are available.

- Liver stiffness by MRE (kPA)
- Liver stiffness by transient elastography (kPA)
- Steatosis by MR-PDFF (percentage)
- Liver metals concentrations obtained from liver biopsy specimen

For endpoints with ordinal outcomes that are only measured at Baseline and Week 48, a Wilcoxon Signed-rank test will be used to compare Baseline versus Week 48 assessment values.

- TSQM-9
- UWDRS

The reference SAS code for a paired sample t-test and a Wilcoxon Signed-rank test can be found in Section 9.5.

For mitochondrial ultrastructure by electron microscopy and Cu deposition in the eye by optical coherence tomography, only by-patient listings will be provided.

7.2.5. Pharmacokinetic and Pharmacodynamic Analyses

For PK and PD endpoints in the Treatment and Extension Periods, analyses will be performed using the respective PK and PD Analysis Sets, or FA Set in the Treatment and Extension Periods.

7.2.5.1. Pharmacokinetic and Pharmacodynamic Sampling

In the Treatment and Extension Periods, plasma samples will be obtained from the blood samples collected pre-dose on the day of all study visits. Patients will be instructed not to take their dose of ALXN1840 on the morning of scheduled study visits, so that PK/PD/biomarker samples are collected at pre-dose trough for measuring total and PUF Mo (PK), total and PUF Cu, cNCC/$NCC_{corrected}$, dNCC, LBC (PD), Cp, and CpC (biomarker). Additional measurements on ALXN1840 PK/PD including Mo and Cu in the tripartite complexes (TPCs) may be conducted.

In the Treatment Period, at the Week 6 (Day 43) and Week 36 (Day 253) Visits, serial blood samples will be collected for plasma PK/PD/biomarker concentrations at the following time points: pre-dose at 0 hour, and at 2- and 4-hours post-dose.

7.2.5.2. BLQ Values / Missing Values

For PK and PD parameter calculations, concentrations below the limit of quantification (BLQ) prior to the first measurable concentration will be set to 0; all other BLQ values will be treated as missing. For PUF Cu parameter calculations, all BLQ values will be set to 0.

For summary statistic calculations, plasma concentration values below the lower limit of quantification (LLOQ) will be set to the LLOQ for analyses using the FA Set and will be set to 0 for analyses using the PK Analysis Set.

For PD and biomarker summary statistic calculations, concentrations below the LLOQ will be set to the LLOQ for analyses using the FA Set and will be set to 0 for analyses using the PD Analysis Set.

7.2.5.3. Pharmacokinetic/Pharmacodynamic Data Presentation Conventions

Pharmacokinetic and PD concentration data will be summarized using the following descriptive statistics: number of patients (N), number of patients with available data (n), arithmetic mean, SD, arithmetic coefficient of variation (arithmetic CV), geometric mean (GM), GMCV, median, minimum, and maximum.

The pre-dose PD concentration will be used as Baseline; the absolute change from Baseline will be calculated as: a post-dose concentration value – Baseline concentration value; and the percent change from Baseline will be calculated as: ([a post-dose concentration value – Baseline concentration value] / Baseline concentration) × 100%.

The following conventions will be applied to PK/PD presentations and summaries.

- For continuous variables, all mean and median values are formatted to 1 more decimal place than the measured value. SD values are formatted to 2 more decimal places than the measured value. Minimum and maximum values are presented with the same number of decimal places as the measured value.
- Date variables are formatted as DDMMMYYYY for presentation. Time is formatted in military time as HH:MM for presentation.

7.2.6. Pharmacokinetic Concentration Analyses

Plasma concentrations of total Mo and PUF Mo (as surrogate measures of ALXN1840 PK) and time data will be presented in a data listing by patient. Plasma concentration data will be summarized separately by analyte, day, and time point using descriptive statistics. When calculating the geometric mean, values of 0 will be discarded.

Mean plasma concentration versus scheduled time profiles will be presented in figures on both linear and semilogarithmic scales. Individual plasma concentration versus actual time profiles will be presented similarly. The time to reach steady state will be graphically assessed by plotting mean plasma trough (pre-dose) concentration observed at the start of the dosing interval (C_{trough}) concentration versus study day in both linear and semilogarithmic scales. Additionally, time to steady state will be evaluated using stepwise testing for linear trend.

7.2.7. Pharmacokinetic Parameter Analyses

The following plasma PK parameters, as data permit, will be calculated for total Mo and PUF Mo on Week 6 (Day 43) and Week 36 (Day 253) using noncompartmental methods with Phoenix® WinNonlin® (Certara USA Inc., Princeton, New Jersey) version 8.0 or higher or SAS® version 9.4 or higher (SAS Institute Inc., Cary, North Carolina), as applicable. Calculations will be based on the actual sampling times relative to the actual reference ALXN1840 dosing times recorded during the study.

- Time delay between the time of dosing and time of appearance of Mo concentration (T_{lag}) in plasma (Week 6 [Day 43] and Week 36 [Day 253])
- Maximum observed concentration (C_{max}) (Week 6 [Day 43] and Week 36 [Day 253])
- Time to maximum concentration (T_{max}) (Week 6 [Day 43] and Week 36 [Day 253])
- Trough (pre-dose) concentration observed at the start of the dosing interval (C_{trough})
- Time of the last quantifiable concentration (T_{last}) (Week 6 [Day 43] and Week 36 [Day 253])
- Observed concentration at the end of the dosing interval (C_τ, where τ = 24 h) (Week 6 [Day 43] and Week 36 [Day 253])

- Area under the plasma concentration versus time curve (AUC) from time 0 to the last quantifiable concentration (AUC_t) (Week 6 [Day 43] and Week 36 [Day 253])
- AUC over the dosing interval (AUC_τ) (Week 6 [Day 43] and Week 36 [Day 253])
- Accumulation ratio (AR) calculated as C_{max}, C_{trough}, and AUC_τ after repeat dosing on Week 36 (Day 253) divided by those after the initial dosing on Week 6 (Day 43):
    - $C_{max,Day253}/C_{max,Day43}$
    - $C_{trough,Day253}/C_{trough,Day43}$
    - $AUC_{\tau,Day253}/AUC_{\tau,Day43}$

Additional plasma PK parameters may be calculated if deemed appropriate.

Pharmacokinetic parameters derived from plasma concentrations of total Mo and PUF Mo will be presented in data listings and summarized by analyte and day using descriptive statistics.

7.2.8. Pharmacodynamic and Biomarker Concentration Analyses

Individual plasma concentration data for measured total Cu, PUF Cu, cNCC/ $NCC_{corrected}$, dNCC, LBC, Cp, and CpC (and Mo and Cu in TPCs if measured) will be summarized separately for each scheduled sampling time using descriptive statistics. When calculating the geometric mean, values of 0 will be discarded. Individual concentration data will be presented in data listings. Individual concentration versus actual time profiles will be provided. Additionally, mean concentration versus nominal time profiles will be provided. Note that the measured, change from Baseline, and percent change from Baseline concentration data for PD and biomarker will be listed, summarized, and presented graphically on linear scales. Similar analysis plan will also be performed for 24-hour urine Cu/Mo/creatinine when applicable.

The dNCC are obtained by inductively coupled plasma mass spectrometry after immunocapture and removal of ceruloplasmin. The daily mean AUEC of dNCC from 0 to 48 weeks, denoted as $AUEC_{0\text{-}48W}$, will be calculated using the linear trapezoidal rule.

$$AUEC_{0-t_n} = \sum_{i=1}^{n} \frac{c_{i-1} + c_i}{2}(t_i - t_{i-1})$$

where, n = number of measured timepoints.

The AUEC is then divided by the number of days to yield $AUEC_{0\text{-}48W}$. $AUEC_{0\text{-}48W}$ of dNCC will be summarized using descriptive statistics including geometric mean and geometric %CV.

In addition to dNCC, cNCC will be calculated by subtracting the amount of Cu bound to Cp from the total plasma Cu level:

$$cNCC[\mu mol/L] = \frac{Total\ plasma\ Cu\ [\mu g/L] - (3.15 \times Cp\ [mg/L])}{63.5\ [\mu g/\mu mol]}$$

For patients dosed with ALXN1840, there is also Cu bound in the tetrathiomolybdate-Cu-albumin TPC, which is addressed by the NCC correction method:

$$NCC_{corrected} = \left(\sqrt{cNCC} - 0.993\sqrt{Mo}\right)^2$$

In the calculation of cNCC and $NCC_{corrected}$, the following adjustments will be made:

- Cu values < LLOQ are set to missing and cNCC will not be derived.
- Cp values < LLOQ are set to 0.
- Mo values < LLOQ are set to 0.
- Negative cNCC values are not utilized and $NCC_{corrected}$ will not be derived.
- $NCC_{corrected}$ will be set to 0 when $0.993\sqrt{Mo} > \sqrt{cNCC}$.

7.2.9. Pharmacodynamic Parameter Analyses

If data permit, the following plasma PD parameters will be calculated in the Treatment Period for the Week 6 (Day 43) and Week 36 (Day 253) Visits using noncompartmental methods for total Cu and PUF Cu (measured and absolute change from Baseline) concentrations with Phoenix® WinNonlin® (Certara USA Inc., Princeton, New Jersey) version 8.0 or higher or SAS® version 9.4 or higher (SAS Institute Inc., Cary, North Carolina), as applicable. Calculations will be based on the actual sampling times relative to the actual reference ALXN1840 dosing times recorded during the study.

- E_{max}: maximum observed PD concentration in plasma
- TE_{max}: time to reach maximum observed PD concentration in plasma
- $AUEC_t$: AUEC from time 0 to last quantifiable concentration time, calculated using the linear trapezoidal rule

Additional PD parameters may be calculated, as necessary.

Parameters for total Cu and PUF Cu in plasma will be summarized separately using descriptive statistics. Geometric mean and geometric %CV will be calculated for AUECs and E_{max}.

7.3. Safety Analyses

All safety analyses will be conducted on the Safety Analysis Set in the Treatment Period. Safety analyses will include all AEs, ECGs, clinical laboratory data, physical examinations, and vital sign measurements using descriptive statistics. By-patient listings will also be provided. No formal hypothesis testing is planned.

The verbatim terms as reported in the CRF by Investigators to identify AEs will be coded using MedDRA version 23.0 or higher and summarized by primary SOC and PT.

7.3.1. Study Duration, Treatment Compliance, and Exposure

Study duration will be summarized for all enrolled patients. Treatment compliance and exposure will be summarized for the Safety Analysis Set.

Study duration is defined as the time from the first dose to the end of the main study period (i.e., Week 48 for patients who do not go into the Extension Period or up to Week 96 for patients in the Extension Period) or study discontinuation date, whichever occurs first.

Compliance to the study treatment regimen will be determined as the number of days the study drug is taken / the number of days the study drug is expected to be taken during active study

treatment × 100. Treatment compliance will be summarized using descriptive statistics. Compliance up to Week 48 will be shown for all patients. Patients who enter the Extension Period will also have their compliance calculated up to Week 96 and only for the duration of the Extension Period from Week 48 to Week 96. If a patient discontinues prematurely from the study, his or her compliance will be based on the period up to the point of discontinuation from the study. Compliance will also be summarized using counts and percentages by compliance category (i.e., ≥ 100% compliance, 80% to < 100% compliance, and 60% to < 80% compliance).

Study drug administration data will be presented in a by-patient listing. The start and end dates for a dose and frequency for each patient will be presented.

Extent of exposure to study drug will be summarized. The duration of exposure will be presented in days and calculated as the last dose of study drug - the first dose of study drug + 1. Duration of exposure, average daily dose (mg), minimum daily dose (mg), and maximum daily dose (mg) will be summarized using descriptive statistics.

7.3.2. Adverse Events

The verbatim terms as reported in the CRF by Investigators to identify AEs will be coded using MedDRA version 23.0 or higher and summarized by primary SOC and PT.

AE toxicity will be evaluated using the NCI CTCAE version 5.0 (published 27 Nov 2017).

AE causality is determined by the Investigator using the following assessment categories: unrelated or related.

TEAEs are defined as those AEs with onset on or after the first dose of treatment (i.e., study drug). Events reported with a partial onset date (e.g., month and year are reported, but the day is missing) will be considered to be treatment emergent if it cannot be confirmed that the event onset was prior to the first dose of study drug based on the available date entries.

7.3.2.1. Overall Summary of Adverse Events

An overall summary of TEAEs will be presented, including the number of events and the number and percentage of patients experiencing AEs. Percentage will be calculated as n / N × 100, where n is the number of patients with events and N is the number of patients in the Safety Analysis Set. The summary will include categories indicating how many events are TEAEs, treatment-emergent SAEs, and treatment-emergent non-SAEs.

A patient-level listing of all TEAEs will be presented. Separate listings will be produced for SAEs, AEs leading to study drug withdrawal, AEs resulting in death, AEs leading to withdrawal from the study, and pre-treatment AEs (PTAEs).

7.3.2.2. Adverse Events and Serious Adverse Events by System Organ Class and Preferred Term

The number of TEAEs and the number and percentage of patients with events will be presented by SOC and PT. Patients are counted once in each SOC and PT. Percentages will be based on the Safety Analysis Set. The SOCs will be listed in descending frequency as will PTs' within each SOC. If needed, PTs will also be ordered alphabetically.

Treatment-emergent SAEs, treatment-emergent non-SAEs, TEAEs leading to withdrawal of study drug, TEAEs leading to death, and PTAEs will be summarized using the same approach.

7.3.2.3. Adverse Events by System Organ Class

The number of TEAEs and the number and percentage of patients with events will be presented by SOC. Patients are counted once in each SOC. Percentages will be based on the Safety Analysis Set.

7.3.2.4. Adverse Events by Preferred Term

The number of TEAEs and the number and percentage of patients with events will be presented by PT. Patients are counted once in each PT. Percentages will be based on the Safety Analysis Set.

7.3.2.5. Adverse Events and Serious Adverse Events by System Organ Class, Preferred Term, and Relationship

The number of TEAEs and the number and percentage of patients with events will be presented by SOC and PT as described in Section 7.3.2.2 by relationship (related and not related). If a patient has > 1 occurrence of an AE, the strongest relationship to study drug will be used in the summary table. If relationship to study drug is missing, the AE will be assumed to be related. A similar analysis will be conducted for treatment-emergent SAEs.

The number of related TEAEs and the number and percentage of patients with related TEAEs will be summarized by SOC and PT, and separately by PT only. The same analyses will be produced for related treatment-emergent SAEs.

Lastly, the number of TEAEs by SOC, PT, and relationship, without taking into account the highest relationship, will be analyzed. A similar analysis will be conducted for treatment-emergent SAEs.

7.3.2.6. Adverse Events and Serious Adverse Events by System Organ Class, Preferred Term, and Toxicity

The number of TEAEs and the number and percentage of patients with events will be presented by SOC, PTs and toxicity (i.e., CTCAE grade). If a patient has > 1 occurrence of an AE, the highest toxicity reported will be used. The number of TEAEs by SOC, PTs and toxicity, without taking into account the highest toxicity, will also be analyzed.

Additionally, a summary of related TEAEs by SOC, PT, and toxicity using the highest toxicity will be presented.

7.3.3. Other Safety

7.3.3.1. Analyses for Laboratory Tests

Protocol-required safety clinical laboratory assessments include tests under Clinical Chemistry, Hematology, Urinalysis, Coagulation, and Other Tests. Actual values and changes from Baseline will be summarized descriptively for patients with available data for each laboratory parameter by analysis visit. Details for Baseline definition and analysis visit window can be referred to in

Appendix 9.4.4. Missing laboratory data will not be imputed. Multiple records may exist for laboratory parameters with the same date and time and when this occurs, the records will be averaged. When multiple records (including averaged records) exist for a same laboratory parameter within the same analysis visit, the record closest to the target date will be used. If there are multiple records with the same distance from the target date, the latest record will be used. A summary for "Last Assessment" will be included for the last available post-baseline result for each patient. A summary of "Worst Post-Baseline" from all post-baseline data will also be included separately for the Treatment Period and for the Extension Period. For categorical test results, actual values will be summarized by analysis visit. All data including that which is only collected at Screening will be included in by-patient data listings. Laboratory measurements will be listed separately by patient, laboratory test, visit and unit.

The Investigator will evaluate any out of normal range laboratory values and make a determination as to whether the observation is not clinically significant (NCS) or clinically significant (CS).

Where applicable, laboratory results will be classified as "low," "normal," or "high" with respect to the parameter-specific reference ranges (i.e., below the lower limit of the normal range, within the normal range, or above the upper limit of the normal range). Contingency tables will be presented for each laboratory parameter to summarize the shift from the Baseline category to all analysis visits. In the tables, count and percentage of patients within each shift category will be included.

Laboratory values outside the normal range will also be summarized and assessed for trends indicating a safety signal. Additionally, a summary and a listing of liver enzyme elevation will be presented.

The following laboratory results will also be graded according the NCI CTCAE version 5.0. Contingency tables will be presented for each laboratory parameter to summarize the shift from the Baseline category to all analysis visits. In the tables, count and percentage of patients within each shift category will be included.

- Hematology: Absolute neutrophil count (neutrophil) (NEUT), total leukocytes (Leukocytes) (WBC), hemoglobin (HGB) (only do the grading for Anemia, no grading for Hemoglobin increased), platelet count (PLAT), Lymphocyte (LYM)
- Coagulation: International normalized ratio (INR)
- Chemistry: Alanine aminotransferase (ALT), albumin (ALB), alkaline phosphatase (ALP), aspartate transaminase (AST), total bilirubin (BILI), creatinine (CREAT), gamma-glutamyl transferase (GGT), glucose (GLUC) (only do the grading the Hypoglycemia, no grading for Hyperglycemia), creatine kinase (CK)

7.3.3.2. Evaluation of Drug Induced Serious Hepatotoxicity (eDISH)

A Hy's law case refers to an increase in aminotransferase > 3x the reference upper limit of normal (ULN), with bilirubin > 2x ULN. Possible Hy's law cases can be visualized with use of Evaluation of Drug Induced Serious Hepatotoxicity (eDISH) plots, a log-log scatter plot where the x-axis is the peak post-baseline ALT as a multiple of ULN, and the y-axis is the peak post-baseline total bilirubin as a multiple of ULN (Guo, 2009).

The following series of figures, adapted from Tesfaldet et al. (2016), will be produced for the Safety Analysis Set.

1. Distribution of ULN of Liver Serum Enzymes. The ULN reference values used in each test (bilirubin, ALT, AST, and ALP) and their respective frequencies in percentiles.
2. Distribution of Baseline Liver Serum Enzymes for each test.
3. Distribution of Baseline Liver Serum Enzymes (xULN), for each test.
4. Status of Baseline Liver Serum Enzymes as Function of ULN. Baseline status categorized as normal, >1x, >1.5x, >2x, >3x, >5x, >10x.
5. eDISH plot for all patients.
6. eDISH plot by quadrant.
7. Panel of eDISH Quadrant Shift Plots at Baseline. Patients are color coded to correspond to the quadrant they belong in the eDISH plot. The data presented in the panel corresponds to each patients' Baseline value as a multiple of the ULN.
8. Panel of Shift Plots by eDISH Quadrants. Patients are color coded to correspond to the quadrant they belong in the eDISH plot. The data is presented as multiples of the Baseline value (BLN) rather than the ULN. The shift in peak post-baseline laboratory value is compared to each patients' Baseline value.
9. Panel of eDISH Shift Plots by eDISH Quadrants. The panels represent the on-treatment eDISH quadrants whereas the colored symbols represent the eDISH quadrants of the Baseline values (not post-baseline) as multiples of ULN.
10. Time Course of Liver Tests as xBLN & xULN.

7.3.3.3. Vital Signs

Actual values and changes from Baseline in vital signs (temperature, systolic and diastolic blood pressure, pulse, and respiratory rate) at each analysis visit will be summarized descriptively. Details for Baseline definition and analysis visit window can be referred to in Appendix 9.4.4. Missing vital signs data will not be imputed and only scheduled assessments will be summarized in tables. When multiple records exist for a same vital sign parameter within the same analysis visit, the record closest to the target date will be used. If there are multiple records with the same distance from the target date, the latest record will be used. A summary for “Last Assessment” will be included for the last available post-baseline result for each patient. Summaries of change from Baseline by category in blood pressure and heart rate will also be created to allow detection of clinically relevant changes; these will be changes of ± 20 mmHg for systolic or diastolic blood pressure and ± 20 BPM for resting heart rate. A by-patient listing of vital signs will be presented, with both scheduled and unscheduled assessments included.

7.3.3.4. Electrocardiogram

Actual values and changes from Baseline in ECG data (heart rate, PR, QRS, QT, QTc, and QTcF intervals) at each visit will be summarized descriptively. Missing ECG data will not be imputed and only scheduled assessments will be summarized in tables.

Single 12-lead ECG results will be classified by the Investigator as "normal," "abnormal, NCS," or "abnormal, CS" and will be summarized in a table. Three-by-three contingency tables will be presented to summarize the shift from the Baseline category to the worst post-baseline value. Summary results will include the count and percentage of patients within each shift category.

The ECG results will also be presented in listings by patient and by visit.

7.3.3.5. Physical Examinations

Results of the abnormal physical examination will be presented in a summary by study visit and body system. Physical examination data (including height and weight) will also be listed at all visits.

7.3.3.6. Other Safety Parameters of Special Interest

No other safety events of special interest have been defined for this study.

7.4. Extension Period Analyses

Actual value and change from Baseline for efficacy endpoints will be analyzed using descriptive summary statistics. Efficacy analyses will be performed using the FA Set in the Extension Period.

Actual value and change from Baseline for safety endpoints will be analyzed using descriptive statistics. Safety analyses will be performed using the Safety Analysis Set in the Extension Period.

All data collected in the Extension Period will be fully listed in the same by-patient listing for the Treatment Period. The same Baseline will be used for both Treatment Period and Extension Period.

7.5. Interim Analyses

Marketing Authorization Application(s) may be submitted before all patients complete the Treatment Periods; therefore, interim analyses of safety data only or safety and efficacy data together may be performed to support these submissions. These analyses will be descriptive only, a full or a subset of the listed outputs will be provided as appropriate, and they will not include formal hypothesis testing and will not be used to adapt the study. These analyses will be performed for the disposition, demographics, Baseline characteristics, exposure, safety, primary, secondary, and exploratory objectives, if sufficient information is available.

8. REFERENCES

Alcorn, J.B. (United Network for Organ Sharing). Letter to: Transplant Professionals. 2015 Nov 23. Available from: https://optn.transplant.hrsa.gov/media/1575/policynotice_20151101.pdf.

Guo T, Senior J, Gelperin K. How a SAS/IntrNet tool was created at the FDA for the detection of potential drug induced liver injury using data with CDISC standard. US FDA, Silver Spring, MD; 2009.

Jayakumar S. Liver transplantation for non-alcoholic fatty liver disease—a review. AME Med J. 2018;3(2):29.

Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology. 2005;41(6):1313-21.

Tesfaldet B, Patel T, Fleischer R, et al. Visualization of liver tests in shift plots integrating multiples of baseline and upper limit of normal. Proceedings of the FDA/CDER/OTS/OCS FDA/CDER/OND/DAVP FDA/NCTR. 2016.

The French Metavir Cooperative Study Group. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. Hepatology. 1994; 20:15-20.

Vallet-Pichard, A., Mallet, V., Nalpas, B., Verkarre, V., Nalpas, A., Dhalluin-Venier, V., Fontaine, H., Pol, S. FIB-4: An Inexpensive and Accurate Marker of Fibrosis in HCV Infection. Comparison with Liver Biopsy and FibroTest. Hepatology 2007; 46(1), 32-36.

9. APPENDICES

9.1. Protocol Schedule of Events

Please refer to Protocol Amendment 5 Section 1.3 Table 1 and Table 2 for schedule of activities.

9.2. Changes From Analyses Specified in the Previous Version of the SAP

Please refer to Section 4.2 in this SAP for changes from analyses specified in the previous version of the SAP.

9.3. Sample Size, Power, and Randomization

Using a t-test statistic, 24 patients in ALXN1840 will provide 92% power at a 1-sided significance level of 0.025 to detect a reduction of 228 µg Cu/g dry liver, with an SD of paired difference of 319, in Cu concentration.

Approximately 28 patients will be enrolled to accommodate a 15% dropout rate, such that approximately 24 evaluable patients complete the study.

9.4. Technical Specifications for Derived Variables

The following derived data will be calculated prior to the analysis.

9.4.1. Adverse Events

The analysis of AEs is described in detail in Section 7.3.2.

The TEAEs are AEs with onset on or after the first study drug dose. If the start date of an AE is partially or completely missing and the end (stop) date and time of the AE do not indicate that it occurred prior to the first dose, then the determination of treatment-emergent status will be based on the following:

- If the start year is after the year of the first study drug dose, then the AE is treatment-emergent; else,
- If the start year is the same as the year of the first study drug dose and
    - The start month is missing, then the AE is treatment-emergent; else if
    - The start month is present and is the same or after the month of the first study drug dose, then the AE is treatment-emergent; else,
- If the start date is completely missing, then the AE is treatment-emergent.

All other AEs are considered pre-treatment AEs.

9.4.1.1. Handling Missing/Partial AE Start Date

The missing or partial AE start date will be imputed for analysis as follows:

- If the year is not missing, but both day and month are missing and
    - If the year of the AE start date is the same as year of the first dose date, then impute AE start date to the first dose date.

  - If the year of the AE start date is not the same as year of the first dose date, then impute the missing day and month of AE start date as June 15.

- If only the day of the AE start date is missing and
  - If the year of the AE start date has the same year and month parts as in the first dose date, then impute the AE start date to the first dose date.
  - If the AE start date has the same year and month parts as in the AE end date, and the day of the AE end date is smaller than 15, then impute the AE start date to the AE end date.
  - Otherwise, impute the missing day of the AE start date as 15.
- If the year is missing, leave the AE start date as missing.

After steps mentioned above, if the imputed AE start date is after the non-missing AE end date, then reset the AE start date to the AE end date.

If imputation results in a date prior to the birth date, then change the AE start date to be equal to the birth date.

If imputation results in a date after the death date, then change the AE start date to be equal to the death date.

Completely missing dates will not be imputed.

9.4.2. Age

Age will be presented as the number of years between the date of birth and the reference date. Age at enrollment will be calculated based on the informed consent date using the following formula:

Age (year) = FLOOR((date of informed consent – date of birth)/365.25)

where FLOOR( ) function returns the integer part of the result.

In cases where only the month and year are provided for a date, the day for the date will be imputed as 15. The missing month will be imputed as June. In cases where the day is observed but the month is missing, the date will be imputed as June 15. In instances when the imputed reference date is earlier than the birth date, the birth date will be used as the reference date.

9.4.3. Analysis Relative Day

Analysis relative day is defined as the actual day relative to the first dose date of study drug. Specifically, the date of interest – the first dose date if the date of interest occurs prior to the first dose date and the date of interest – the first dose date + 1 if the date of interest occurs on or after the first dose date.

9.4.4. Analysis Visit Windowing

In the analysis of data summarized by visit, all data collection will be reassigned an analysis visit where data are scheduled for collection based on the actual days relative to Baseline. All visits will be assigned a target study day. For the determination of target days, weeks will be assumed to have 7 days. Baseline will have a target study day of 1. Thus, following Week x would have a

target day of x×7+1. For each assessment, the post-baseline period will be divided up using the scheduled visit's target days. The lower bound of each visit interval will be calculated as the mid-point between the target day and the previous visit's target day in the following manner: study day interval lower bound = target study day - ((target study day - last target study day) / 2). Since each assessment has its own schedule, the study day interval needs to be defined separately according to the SoA listed in Section 9.1. See below Table 2 for an example of Vital Signs.

Table 2: Analysis Visit Windows (Example of Vital Signs)

Scheduled Visit Week	Target Study Day	Study Day Interval
Baseline	1	The last measurement on or prior to the first dose of study drug
Week 4	29	2 to 35
Week 6	43	36 to 49
Week 8	57	50 to 70
Week 12	85	71 to 105
Week 18	127	106 to 147
Week 24	169	148 to 210
Week 36	253	211 to 294
Week 48 (without Extension)	337	295 to 350
EOS (without Extension)	365	351 to the last date of patients don't enter Extension
Week 48 (with Extension)	337	295 to 378
Week 60	421	379 to 462
Week 72	505	463 to 546
Week 84	589	547 to 630
EOS (with Extension)	673	631 to the last date of patients enter Extension

Abbreviations: EOS = end of study

9.4.5. Baseline

Baseline is defined as the last assessment yielding non-missing valid values taken prior to the first dose of study drug. The same Baseline will be used for both Treatment Period and Extension Period.

9.4.6. Change From Baseline

Change from Baseline will be calculated as post-baseline value - Baseline value when both are not missing. If either value is missing, change from Baseline will be set to missing. The variable is for post-baseline visits only.

9.4.7. Body Mass Index

The BMI is derived as follows: weight (kg) / [height (cm) / 100]^2

9.4.8. Duration

Duration variables will be calculated in days as the end date – the start date + 1, unless otherwise specified.

Disease duration will be presented as the number of years between the date of the first infusion and the date of diagnosis.

9.4.9. Medications and Therapies

Concomitant medications/therapies are any medications/therapies with administration dates and times on or after the date and time of the first study drug dose. If the start date of a medication or therapy is partially or completely missing and the end (stop) date and time of the medication/therapy do not indicate that it occurred prior to the first dose, then the determination of concomitant status will be based on the following:

- If the start year is after the year of the first study drug dose, then the medication/therapy is concomitant; else,
- If the start year is the same as the year of the first study drug dose; and
  - The start month is missing, then the medication/therapy is concomitant; else if
  - The start month is present and is the same or after the month of the first study drug dose, then the medication/therapy is concomitant; else,
- If the start date is completely missing, then the medication/therapy is concomitant.

All other medications/therapies are considered prior medications/therapies and could occur from the 30 days prior to informed consent up to the Screening Period and prior to the first dose of study drug.

9.4.10. MELD Score

The MELD score uses the patient's values for serum bilirubin, serum creatinine, and the INR to stratify severity of end-stage liver disease and was developed for transplant planning. The initial MELD score, $MELD_{(i)}$, is calculated according to the following formula:

$$MELD_{(i)} = 3.78 \times \ln[\text{serum bilirubin (mg/dL)}] + 11.2 \times \ln[\text{INR}] + 9.57 \times \ln[\text{serum creatinine (mg/dL)}] + 6.43$$

Creatinine, bilirubin, and INR values < 1.0 are set to 1.0, and creatinine values > 4.0 are set to 4.0 when calculating $MELD_{(i)}$. Additionally, creatinine, bilirubin, and INR are rounded to the tenth decimal place prior to performing the calculation. The initial MELD score is then rounded to the nearest integer. The maximum MELD score is 40.

A modification to the MELD score exists for MELD scores > 11 (Alcorn, 2015) and this will be applied to remain up to date with the current guidance. Thus, if MELD(i) is > 11, the MELD score is recalculated as follows:

$$MELD = MELD_{(i)} + 1.32 \times (137 - \text{sodium (mmol/L)}) - 0.033 \times MELD_{(i)} \times (137 - \text{sodium (mmol/L)})$$

Sodium values < 125 mmol/L will be set to 125, and values > 137 mmol/L will be set to 137.

The MELD score will be calculated each time the appropriate clinical chemistry parameters are obtained at visits listed in Section 9.1. If any of the parameters required to compute the MELD score are unavailable at any given visit, the MELD score will not be derived for that patient visit.

9.4.11. Modified Nazer Score

The modified Nazer score will be calculated by a central laboratory at visits listed in Section 9.1. If any of the 5 parameters required to compute the modified Nazer score are unavailable at any given visit, the modified Nazer score will be set to missing for that patient visit.

The score for an individual analyte (bilirubin, AST, INR, leukocytes, and albumin) should be derived from Table 3, and then, all 5 scores will be added to get the final score.

Table 3: Modified Nazer Score

Score	Bilirubin (µmol/L)	AST (IU/L)	INR	Leukocytes (10^9/L)	Albumin (g/L)
0	0-100	0-100	0-1.29	0-6.7	>45
1	101-150	101-150	1.3-1.6	6.8-8.3	34-44
2	151-200	151-300	1.7-1.9	8.4-10.3	25-33
3	201-300	301-400	2.0-2.4	10.4-15.3	21-24
4	> 301	> 401	> 2.5	> 15.4	< 20

Abbreviation: AST = aspartate transaminase; INR = international normalized ratio.

9.4.12. CLDQ

The CLDQ is a 29-item instrument with 6 domains (Fatigue, Activity, Emotional function, Abdominal symptoms, Systemic symptoms, Worry) that measures longitudinal change over time in patients with chronic liver disease. Each domain is scored from 1 (worst possible function) to 7 (best possible function). Scoring of the questionnaire is performed by dividing each domain score by the number of items per domain. Overall CLDQ score is obtained by adding scores for each item and dividing by the total number of items.

9.4.13. TSQM-9

The TSQM has 9 questions making up 3 components: effectiveness score, convenience score, and global satisfaction score. Details on deriving the component scores are given below, including how to handle a single missing question within a component. If > 1 question is missing for a component, the component score cannot be derived.

The effectiveness score is composed of Questions 1 to 3, each of them ranging from 1 to 7. The score is calculated as follows: (Q1+Q2+Q3-3)/18×100. If one of the questions is missing, the score is calculated as follows (including the missing item): (Q1+Q2+Q3-2)/12×100.

The convenience score is composed of Questions 4 to 6, each ranging from 1 to 7. The score is calculated as follows: (Q4+Q5+Q6-3)/18×100. If one of the questions is missing, the score is calculated as follows (including the missing item): (Q4+Q5+Q6-2)/12×100.

The global satisfaction score is composed of Questions 7 to 9, where Questions 7 and 8 range from 1 to 5 and Question 9 ranges from 1 to 7. The score is calculated as follows: (Q7+Q8+Q9-3)/14×100. If either Question 7 or 8 is missing, the score is calculated as follows (including the missing item): (Q7+Q8+Q9-2)/10×100. If Question 9 is missing, the score is calculated as follows: (Q7+Q8-2)/8×100.

9.4.14. UWDRS

The algorithms for calculating the UWDRS subscores and subscales are given below. The UWDRS total score will be the sum of the 3 subscores.

UWDRS I: Consciousness; Maximum score of 3

- Set to Question 1, range 0 to 3.

UWDRS II: Disability; Maximum score of 40

- Calculate the sum of Question 2 to Question 11, each question has range 0 to 4
- If all 10 items are populated, the subscore will be the sum calculated above. If 8 or 9 items are populated, pro-rate the score by dividing by the number of answered items and multiplying by 10. This is done in order to estimate the value the patient could have achieved if they had answered all the questions. If < 8 items are populated, the score will not be derived.
    - For example, a patient has answered 8 of 10 questions and the total of his or her answered questions is 23. The pro-rated score will then be calculated as follows: 23/8×10 = 28.75.

UWDRS III: Neurological Status; Maximum score of 175

- Calculate the following new scores for Questions 12 to 34:
    - Question 12: range 0 to 4

      Set to Q12A
    - Question 13: range 0 to 6

      If Q13=0 then set to 0;
      else if Q13=1 then do;

      if Q13A>2 then set to Q13A
      else set to Q13A+Q13B
      end;
    - Question 14: range 0 to 1

      Set to Question 14
    - Question 15: range 0 to 16

      Set to Q15A+Q15B+Q15C+Q15D
    - Question 16: range 0 to 4

      Set to Q16

- Question 17: range 0 to 20

  Set to Q17A+Q17B+Q17C+Q17D+Q17E

- Question 18: range 0 to 8

  Set to Q18A+Q18B

- Question 19: range 0 to 8

  Set to Q19A+Q19B

- Question 20: range 0 to 4

  Set to Q20

- Question 21: range 0 to 16

  Set to Q21A1+Q21A2+Q21B1+Q21B2

- Question 22: range 0 to 8

  Set to Q22A+Q22B

- Question 23: range 0 to 8

  Set to Q23A+Q23B

- Question 24: range 0 to 8

  Set to Q24A+Q24B

- Question 25: range 0 to 4

  Set to Q25

- Question 26: range 0 to 8

  Set to Q26A+Q26B

- Question 27: range 0 to 4

  Set to Q27

- Question 28: range 0 to 10

  If Q28=0 then set to 0;
  else if Q28=1 then do;
  if Q28A>2 then set to Q28A
  else set to Q28A+Q28B+QS28C
  end;

- Question 29: range 0 to 10

  If Q29=0 then set to 0;

  else if Q29=1 then do;

  if Q29A1+Q29A2>2 then set to Q29A1+Q29A2
  else set to Q29A1+Q29A2+Q29B+Q29C
  end;

- Question 30: range 0 to 24

  Set to Q30A+Q30B+Q30C+Q30D+Q30E+Q30F
- Question 31: range 0 to 1

  Set to Q31
- Question 32: range 0 to 1

  Set to Q32
- Question 33: range 0 to 1

  Set to Q33
- Question 34: range 0 to 1

  Set to Q34

- Add up the newly calculated scores for Questions 12 to 34. If all 23 items are completely populated, the subscore will be that sum. If 20 to 22 items are fully populated, then calculate the maximum score of the answered items. Next, divide the sum of the responses to Questions 12 to 34 by the maximum score and multiply by 175 to obtain the subscore. This is done in order to estimate the value the patient could have achieved if they had answered all the questions. If < 20 items are fully populated, then do not derive. If part of a multi-component question is answered, the maximum score of the answered components will be used to derive the subscore. However, partially answered questions will not be considered as fully populated (and therefore, will be viewed as a null response) when counting the number of fully answered questions in order to obtain a minimum of 20 fully answered questions. The exception to this latter rule is for Questions 13, 28, and 29. If the > 2 condition is met with populated components, the question is considered to be fully populated regardless if the unused components are null.
  - For example, a patient has responded to 21 questions and left Questions 21 and 22 null. The total of their responses to the answered questions is 97 and the maximum score of the answered questions is 151. The pro-rated score will then be calculated as: 97/151×175=112.41721854.
  - For example, a patient has responded to 22 questions and part of Question 17 so they will be considered as answering 22 questions. The total of their response to the answered questions is 111 and the maximum score of the answered questions and answered components (including any partially answered questions) is 167. The pro-rated score will then be calculated as 111/167×175=116.317365.

UWDRS II Activities of Daily Living (ADL): Mobility, Feeding, Dressing, Grooming, Taking a Bath or Shower, and Toilet Use; Maximum score of 10

- Take average of the transformed items for Q2, Q7, to Q11

UWDRS II Individual symptoms:

- Salivation: Q5, range 0 to 4
- Swallowing: Q6, range 0 to 4

UWDRS III Tremor: Tremor at Rest, Head Tremor, Tremor in Arms, Postural Tremor Lower Extremity, and Finger-to-Nose Test; Maximum score of 10

- Take average of the transformed items for Q15, Q16, Q21, Q22, Q24

UWDRS III Dystonia: Speech, Facial Expression, Rigidity, Finger Taps, Rapid Alternating Hand Movements, Handwriting, Leg Agility, Cervical Dystonia, Arm and Hand Dystonia, Arising from Chair, Posture, and Gait; Maximum score of 10

- Take average of the transformed items for Q12, Q13, Q17 to Q20, Q23, and Q25 to Q29.

UWDRS III Individual symptoms

- Speech: Q12, range 0 to 4.
- Handwriting: Q20, range 0 to 4.

For each new subscale (ADL, Tremor, Dystonia), at least 50% of items need to be answered in order to calculate new subscale. In order to standardize responses, these items will be transformed to a 0 to 10 scale and then take the average of the transformed items.

- For single-item score, divide the score by the range of options (0-4, then divide by 4) to transform it to 0-1 and then multiplied by 10.
- For multi-item score, prorate on the number of sub-items been answered (Q15, Q17, Q18, Q19, Q21, Q22, Q23, Q24, and Q26).
    - For example, Question 15 is a multi-component item – Tremor at rest. There are 4 components: 1) Left Upper Extremity; 2) Right Upper Extremity;3) Left Lower Extremity; and 4) Right Lower Extremity. These 4 components were summed to create “Item 15”, which had a possible range of “0-16” If < 4 sub-component scores were available, use the possible range for the number of items that were answered (e.g., if 3 components were available, then the range was 12).
    - For item 13 (maximum score =6), item 28 (maximum score=10) and item 29 (maximum10), use fixed maximum score to standardize the score.

For the individual symptom scale (salivation, swallowing, speech, and handwriting), the original scale (0-4) will be used.

9.5. Additional Details on Statistical Methods

The reference SAS codes used for efficacy analyses are as follows:

Multiple Imputation

Step 1:

```
proc mi data=xxx nimpute=20 seed=54321 out=yyy;
   mcmc impute=monotone;
   var BASE W48;
run;
```

Step 2:

```
proc ttest data=xxx;
   by _imputation_;
   paired BASE*W48;
run;
```

Step 3:

```
proc mianalyze data=xxx;
   by AVISIT;
   modeleffects ESTIMATE;
   stderr STDERR;
run;
```

Paired sample t-test

```
proc ttest data=xxx;
   paired BASE*W48;
run;
```

Wilcoxon Signed-rank test

```
proc univariate data=xxx;
   var CHG;
run;
```

MMRM

```
proc mixed data=xxx method=reml;
   class AVISITN USUBJID;
   model CHG=AVISITN BASE / solution ddfm=kr cl;
   repeated AVISITN / subject=USUBJID type=UN;
   lsmeans AVISITN / cl;
run;
```