Statistical Analysis Plan Amendment 4

Study ID: 212620

Official Title of Study: A two-part, randomized, placebo controlled, double blind, multicenter, Phase 3 study to evaluate the efficacy and safety of linerixibat for the treatment of cholestatic pruritus in participants with primary biliary cholangitis (PBC).

NCT ID: NCT04950127

Date of Document: 08 Nov 2024

Information Type:	Statistical Analysis Plan (SAP)

TITLE PAGE

Protocol Title:	A two-part, randomized, placebo controlled, double blind, multicenter, Phase 3 study to evaluate the efficacy and safety of linerixibat for the treatment of cholestatic pruritus in participants with primary biliary cholangitis (PBC).
Study Number:	212620
Compound Number:	GSK2330672 (linerixibat)
Abbreviated Title:	*G*lobal *L*inerixibat *I*tch *ST*udy of *E*fficacy and Safety i*N* PBC
[Acronym:]	GLISTEN

Sponsor Name: GlaxoSmithKline Research & Development Limited

Regulatory Agency Identifier Number(s)

IND: 130391

EudraCT: 2021-000007-21

Registry	**ID**
Clintrials.gov	*NCT04950127*

TABLE OF CONTENTS

PAGE

LIST OF TABLES

PAGE

LIST OF FIGURES

PAGE

VERSION HISTORY

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
SAP	01 November 2021	Protocol amendment [03] (02-Sep-2021)	Not Applicable	Original version
SAP Amendment 1	28 June 2022	Protocol amendment [03] (02-Sep-2021)	Added non-Metasite ITT analysis set	In line with regulatory feedback
			Text added to clarify how LS means are averaged in MMRM analysis	In line with regulatory feedback
			Added MMRM analysis of non-Metasite ITT population for the primary endpoint	In line with regulatory feedback
			Clarified the definition of COVID 19 related intercurrent events and that hypothetical strategy will not be applied to supportive secondary endpoints	In line with regulatory feedback
			Added empirical cumulative distribution function (eCDF) graph for primary, week 2, and responder endpoints.	In line with regulatory feedback
			Added exposure adjusted incidence rate for safety analyses	In line with regulatory feedback
			Added analyses of exploratory endpoints, including PK for China, Japan and East Asia subpopulations	Not previously addressed

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
			Added baseline bile acid binding resin (BABR) covariate	Added for biomarker exploratory analysis
			Added study population analyses	Not previously addressed
			Added data derivation rule	Not previously addressed
SAP Amendment 2	19 April 2024	Protocol amendment [04] (20-Nov-2023)	Updated the baseline definition for safety assessments to be the last assessment prior to first dose, instead of Visit 3.	In line with regulatory feedback
			Updated baseline definition for biomarkers to be the average of screening 2 and last assessment prior to first dose (or either value in cases where only one of these values is present).	Not previously addressed
			Updated list of regions in Covariate/Subgroup analyses Section.	Not previously addressed
			Added the predictive probability of success and the conditional power under the Interim Efficacy Analysis Section.	To align with IDMC Charter
			Clarified the calculations of exposure	In line with regulatory feedback

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
			for the exposure adjusted incidence rate.	
			Removed empirical cumulative distribution function (eCDF) graph for change from baseline MWDI score at week 2 as it is not required. Additional eCDF graphs for the other endpoints added, using imputed data.	Not previously addressed
			Mexico specific analysis removed.	No longer required
			In the Subpopulation Analyses Section; defined a Non-BABR ITT East Asia population and clarified the definition of the China and East Asia populations.	Not previously addressed
			Renamed the primary interim analysis set and clarified in the definition that it will include approximately 100 subjects, in line with the study protocol. Also updated the definition of the interim safety population to be analyzed according to treatment as randomized, rather than actually received.	Not previously addressed
			Updated Prior and Concomitant Medications Section to specify that the WHO	Not previously addressed

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
			drug dictionary will be used for all summaries, and summaries will be reported by ATC classification.	
			Updates to Analysis Window Section for unscheduled and withdrawal visits.	In line with regulatory feedback
			Removed RoW Metasite from Study Design Section since no RoW metasites will be included in this study.	No longer required
			Clarified estimands for safety endpoints and added a supplementary safety estimand.	Not previously addressed
			Updated Handling of Partial Dates section, to allow imputation of partial dates when calculating disease duration.	Not previously addressed
			Clarified definitions of treatment states for AEs, and intervention emergent flag for AEs.	Not previously addressed
			Clarified derivation of Adverse Events of Special Interest (AESI).	Not previously addressed
			Updated baseline definition for Laboratory data to consider cases where only 1 sample is available. Added criteria for handling values below LOQ.	Not previously addressed

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
			Clarified the list of Liver related SMQs.	Not previously addressed
			Added calculation of exposure adjusted incidence rate for Supplementary Safety Estimand (while-on-treatment strategy) to Section 6.3.8.2.	Not previously addressed
			Derivation of baseline concomitant medication variables clarified in Section 6.3.10.	Not previously addressed
			Clarified the derivations of Intercurrent Events in Section 6.3.11.	Not previously addressed
			The terminology of Mean Worst Daily Itch (MWDI) Score was changed to Weekly Itch Score (WIS). The calculation remains the same.	In line with Protocol amendment 04
			Clarification of previously omitted Part B endpoints: In exploratory endpoints, added the Quality of Life and Patient-Reported Outcomes (PROs) endpoints for Part B on MSS, MFS, ESS, BDI-II, PGI-S, and PGI-C. Clarified that exploratory biomarker endpoints and the effect of bile acid binding	In line with Protocol amendment 04

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
			resins on linerixibat pharmacodynamics will be assessed in Part B. In exploratory markers of liver fibrosis endpoints, clarified that ELF will be assessed in Part B.	
			Additional details added to Sensitivity Analysis to clarify multiple imputation model.	Clarification of imputation model
			Added details for the new statistical analysis strategy involving a primary analysis after all participants have completed the Part A intervention period and a final analysis after all participants have completed the Part B intervention period, and the follow-up period.	In line with Protocol amendment 04
			Additional e-DISH plots added.	Not previously addressed
			Added ranges for Chronic Kidney Disease Stage, for use in PK summaries.	Not previously addressed
			New analysis added for PGI-C, where the percentage of participants with improved PGI-C will be reported.	Not previously addressed

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
			Additional analysis of ALP and Bilirubin added for participants with ALP ≥1.67x ULN at baseline.	Not previously addressed
			Updated the analyses required for each subpopulation (China, Japan, East Asia).	To align with requirements for each region
			Additional covariates added to Section 4.6.1. And clarified which covariates are to be included in the subgroup analysis.	Not previously addressed
			Clarified handing of participants in the Safety population who receive incorrect treatment (Section 3).	Not previously addressed
			Clarified that AE/SAEs starting in Part A that last to Part B will be counted in Part A only.	Correction due to how the data is recorded
			Removed table for pre-treatment SAEs, these will be reported in listing only. Also removed table for post-treatment SAEs, since these will now be captured in the supplementary safety estimand.	Not previously addressed
			Criteria for Potential Clinical Importance updated (Section 6.3.1).	Not previously addressed

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
SAP Amendment 3	05 August 2024	Protocol amendment [04] (20-Nov-2023)	Included IL-31 biomarker in list of biomarkers to be analysed.	Clarification that IL-31 will also be analyzed as part of the exploratory biomarker objective
			Clarified that ITT population for Part B will be all participants randomized who do not withdraw in Part A. And the Safety population for Part B will be all subjects who take a least one dose in Part B.	Not previously addressed
			Removed Child Pugh derivation since it is not possible to derive at baseline for all subjects.	Not previously addressed
			Clarified that all AEs will be reported using the supplementary estimand.	Not previously addressed
			Added explanation regarding process for de-identification of participant data for public disclosure (Section 6.1.2)	Not previously addressed
			Added a new section for Electronic Patient-reported Outcome (ePRO) Compliance.	Based on updated internal guidance
			Clarified derivation of study period.	Not previously addressed

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
SAP Amendment 4	08 Nov 2024	Protocol amendment [04] (20-Nov-2023)	Subgroup analyses by sex, age, race, and ethnicity will be performed for the primary efficacy endpoint, and for common AEs.	In line with FDA feedback
			Specified the planned shift parameters for the tipping point analysis.	In line with FDA feedback
			Additional details provided for shrinkage estimation.	In line with FDA feedback
			Clarified multiple imputation model.	In line with FDA feedback
			Clarified that MedDRA and WHO dictionaries available at the time of the primary analysis will be used for the final analysis also.	Not previously addressed
			Clarified Part B treatment start date for cases of treatment switch on the same day.	Not previously addressed

1. INTRODUCTION

The purpose of this SAP is to describe the planned analyses to be included in the Clinical Study Report (CSR) for Study 212620. Details of the planned interim analysis, as well as the final analyses, are provided.

This Phase 3 study will investigate the efficacy and safety of 24 weeks of 40 mg BID oral linerixibat, for the treatment of pruritus in PBC patients with moderate to severe cholestatic pruritus (Part A), as well as assess the return of itch over 8 weeks after withdrawal of 40 mg BID oral linerixibat compared to placebo (Part B).

Participants will record itch scores twice daily using an electronic Diary (eDiary). A 0-10 Numerical Rating Scale (NRS), which has been validated in the PBC population, will be used to assess itch with a response option of 0 representing no itching and 10 representing the worst imaginable itching. Every morning the participant will record the worst itch experienced the night before and every evening, the participant will record the worst itch experienced that day.

The itch scores for any day, week, and month will then be determined as follows:

- **Worst Daily Itch (WDI) score:** the worst of the two scores recorded daily will be considered the score for that day
- **Weekly Itch Score (WIS) (formerly referred to as the Mean Worst Daily Itch [MWDI]):** the average of the WDI scores in one week
- **Monthly Itch Score (MIS):** the worst WIS for that month (i.e., worst week score of the 4 weeks)

Participant's sleep will also be evaluated using the daily eDiary. Participants will record a sleep interference score each morning using a 0-10 NRS, where 0 represents no sleep interference and 10 represents complete sleep interference. For each week, the average of the **Daily Sleep Scores (DSS)** will be used to calculate the **Weekly Sleep Score (WSS)**. The **Monthly Sleep Score (MSS)** is the worst WSS for that month (i.e., worst week score of the 4 weeks).

The Primary Biliary Cholangitis-40 Questionnaire (PBC-40) is a disease-specific health-related QoL instrument consisting of 6 domains: social, emotional, symptoms, fatigue, itch (including an item on sleep disturbance from itching), and cognitive. This study will further evaluate the effect of linerixibat on the PBC-40 domains.

1.1. Objectives, Estimands and Endpoints

Objectives*	Endpoints
Primary	
To investigate the effect of treatment with oral linerixibat compared with placebo on itch in PBC patients with cholestatic pruritus over 24 weeks (Part A)	• Change from Baseline in Monthly Itch Scores over 24 weeks using a 0-10 numerical rating scale (NRS)
Secondary	
To evaluate the early effects of oral linerixibat compared to placebo on itch in PBC patients with cholestatic pruritus (Part A)	• Change from baseline in Weekly Itch score at Week 2
To characterize the effects of treatment with oral linerixibat compared with placebo on health related QoL (Part A)	• Change from Baseline in Monthly Sleep Score as measured by 0-10 NRS over 24 weeks • Change from Baseline in PBC-40 domain scores at Week 24
To evaluate the effects of 24 weeks of treatment with oral linerixibat compared to placebo on itch response rates in PBC patients with cholestatic pruritus (Part A)	• Responder defined as achieving a ≥2-point reduction from Baseline in the Monthly Itch score at Week 24. • Responder defined as achieving a ≥3-point reduction from Baseline in the Monthly Itch score at Week 24. • Responder defined as achieving a ≥4-point reduction from Baseline in the Monthly Itch score at Week 24.
To investigate the treatment effect of oral linerixibat compared with placebo on Patient's Global Impression of Severity (PGI-S) and Patient's Global Impression of Change (PGI-C) throughout the treatment period (Part A)	• Change from baseline in Patient's Global Impression of Severity (PGI-S) over 24 weeks • PGI-C over 24 weeks
To evaluate the effects of treatment with linerixibat on markers of PBC disease activity and progression (Part A)	• Change from baseline in ALP at Week 24 • Change from baseline in bilirubin at Week 24

Objectives*	Endpoints
Safety	
• To evaluate the safety of oral linerixibat compared with placebo (Part A and Part B)	Clinical assessments including, but not limited to: • Adverse Events (AEs) and Serious Adverse Events (SAEs) • Vital signs • 12-lead Electrocardiogram (ECG) • Clinical laboratory evaluation (including liver chemistry panel and fasting lipids)
Exploratory	
• To evaluate the gastrointestinal (GI) tolerability of oral linerixibat compared with placebo on gastrointestinal symptoms (Part A and Part B)	• Change from Baseline in Gastrointestinal Symptom Rating Scale (GSRS) over time
• To evaluate the effects of oral linerixibat compared to placebo on the percentage of itch response days (Part A)	• Percentage of days participant achieves a ≥2-point reduction from Baseline in the Worst Daily Itch score over 24 weeks. • Percentage of days participant achieves a ≥3-point reduction from Baseline in the Worst Daily Itch score over 24 weeks. • Percentage of days participant achieves a ≥4-point reduction from Baseline in the Worst Daily Itch score over 24 weeks.
• To investigate the improvement, maintenance, or return of itch (Part B)	• Change from Baseline (Part A) in Weekly Itch Score at 8 weeks • Change from Baseline (Part B) in Monthly Itch Score over 8 weeks (maintenance)

• To investigate the improvement, maintenance, or decline of health-related quality of life (QoL) and other patient-reported outcomes (Part B)	• Change from Baseline in
	  • PBC-40 domain scores at 8 weeks
	  • Monthly Sleep Score at 8 weeks
	  • Monthly Fatigue Score at 8 weeks
	  • Epworth Sleepiness Scale at 8 weeks

Objectives*	Endpoints
	• BDI-II at 8 weeks • PGI-S at 8 weeks • PGI-C at 8 weeks
• To evaluate the effects of oral linerixibat compared with placebo on exploratory biomarkers (Part A and Part B)	• Change from Baseline in concentrations of exploratory biomarkers including[#]: • Serum C4 • Total serum bile acids • Autotaxin • Fibroblast Growth Factor-19 (FGF-19)
• Explore the effect of bile acid binding resins on linerixibat Pharmacodynamics (PD) (Part A and Part B)	• Linerixibat serum biomarkers including[#]: • Serum C4 • Total serum bile acids • Autotaxin, • FGF-19
• To characterize the PK in PBC participants (Part A)	• Determination of PK parameters by population PK analysis. [†]
• To evaluate the effect of linerixibat compared to placebo on non-invasive markers of liver fibrosis (where available) (Part A and Part B, if applicable)	• Change from Baseline (Part A) in Enhanced Liver Fibrosis (ELF) at Week 24 • Change from Baseline (Part B) in ELF at week 8 • Change from Baseline in Transient Elastography (Fibroscan) at Week 24
• To further characterize the effects of oral linerixibat compared with placebo on symptoms and health-related quality of life (Part A)	• Change from Baseline in Epworth Sleepiness Scale (ESS) over 24 weeks • Change from Baseline in Beck Depression Inventory-II (BDI-II) at 24 weeks • Change from Baseline in Monthly Fatigue Score as measured by 0 – 10 NRS over 24 weeks.

*Linerixibat refers to linerixibat with stable background itch therapy if applicable. Placebo refers to placebo with stable background itch therapy if applicable.

[#] IL-31, a cytokine recently proposed to be associated with pruritus in PBC will also be included.

[†]See Section 4.8 which provides the rationale for changes related to this endpoint.

Estimands for Primary and Secondary Objectives:

Unless otherwise specified, all primary and key secondary (as detailed in Section 2.2) study objectives will be assessed using estimands defined with the following common elements in terms of population, treatment comparison and approach for managing intercurrent events.

	Primary and key Secondary Endpoints	Supportive Secondary Endpoints
Population	• PBC patients with cholestatic pruritus	
Treatment	• Linerixibat + background itch therapy if applicable • Placebo + background itch therapy if applicable	
Intercurrent events:		
Permanent treatment discontinuation, disruptions in treatment or treatment delays unrelated to the COVID-19 pandemic	Addressed with treatment policy strategy, i.e., regardless of the intercurrent event occurring.	Addressed with treatment policy strategy, i.e., regardless of the intercurrent event occurring.
Permanent treatment discontinuation, disruptions in treatment, or treatment delays related to the COVID-19 pandemic	Addressed with hypothetical strategy, i.e., the outcomes impacted by the COVID-19 pandemic related intercurrent events will be discarded.	Addressed with treatment policy strategy, i.e., regardless of the intercurrent event occurring.
Change in background itch therapy* or use of rescue medication	Addressed with treatment policy strategy, i.e., regardless of the intercurrent event occurring.	Addressed with treatment policy strategy, i.e., regardless of the intercurrent event occurring.

*Change in background itch therapy is defined as any change in dose of background therapy, addition of a new background therapy, or stopping an existing background therapy.

For supportive secondary endpoints (as defined in Section 2.2), treatment policy will be used for all the intercurrent events.

Summary Measures for the study objectives are as follows:

Endpoint	Summary Measure
Primary objective	
Change from Baseline in Monthly Itch Scores over 24 weeks using a 0-10 NRS	Difference in means averaged over 24 weeks between treatment groups
Secondary objectives	
Change from baseline in Weekly Itch Score at week 2	Difference in means between treatment groups
Change from Baseline in Monthly Sleep Score as measured by 0-10 NRS over 24 weeks	Difference in means averaged over 24 weeks between treatment groups
Responder defined as achieving a ≥2-point reduction from Baseline in the Monthly Itch score at Week 24	Difference in proportions between treatment groups
Responder defined as achieving a ≥3-point reduction from Baseline in the Monthly Itch score at Week 24	Difference in proportions between treatment groups
Responder defined as achieving a ≥4-point reduction from Baseline in the Monthly Itch score at Week 24	Difference in proportions between treatment groups
Change from Baseline in PBC-40 domain scores at Week 24	Difference in means between treatment groups
Change from baseline in PGI-S over 24 weeks	Difference in means averaged over 24 weeks between treatment groups
PGI-C over 24 weeks	Difference in means averaged over 24 weeks between treatment groups
Change from baseline in ALP at Week 24	Difference in means between treatment groups
Change from baseline in bilirubin at Week 24	Difference in means between treatment groups

Rationale for estimand: Interest lies in the treatment effect when medication is taken for the entire study duration. For participants discontinuing randomized medication, having disruptions in treatment or treatment delays unrelated to the COVID-19 pandemic, changing background itch therapy, or using rescue medication, the use of a treatment policy strategy recognises that this could be due to an unfavourable cause. The study is designed to evaluate treatment effects attributable to linerixibat and it is important that this effect is not confounded by site closure or lockdown caused by the COVID-19 pandemic. Therefore, for participants who have permanent treatment discontinuation,

disruptions in treatment, or treatment delays related to the COVID-19 pandemic (see Section 4.2.1 for definition), a hypothetical strategy will be used. The definition of COVID-19 pandemic related intercurrent events and the corresponding strategy are subject to change to be consistent with regulatory agencies' evolving recommendations.

Supplementary Estimand for the Primary Objective

A supplementary estimand for the primary objective will be defined to assess the hypothetical treatment effect of linerixibat compared to placebo in the absence of intercurrent events including treatment discontinuation, disruptions in treatment or treatment delays, changes in itch therapy or use of rescue. In this case, all daily itch score data impacted by the intercurrent events will be discarded.

Estimands for Safety Objective:

Unless otherwise specified, Safety endpoints will be assessed using estimands defined with the following common elements in terms of population, treatment comparison and approach for managing intercurrent events.

	Safety Endpoints Including: • **Adverse Events (AEs) and Serious Adverse Events (SAEs)**	**Safety Endpoints Including:** • **Vital signs** • **12-lead ECG** • **Clinical laboratory evaluation (including liver chemistry panel and fasting lipids)**
Population	• PBC patients with cholestatic pruritus	
Treatment	• Linerixibat + background itch therapy if applicable • Placebo + background itch therapy if applicable	
Intercurrent events:		
Permanent treatment discontinuation due to any reason	Addressed with while on-treatment strategy, i.e., considering data prior to intercurrent event. This strategy is to summarise the events while the participants are exposed to the treatment.	Addressed with treatment policy strategy, i.e., regardless of the intercurrent event occurring
Disruptions in treatment or treatment delays due to any reason (related or unrelated to the	Addressed with treatment policy strategy, i.e., regardless of the intercurrent event occurring	Addressed with treatment policy strategy, i.e., regardless of the intercurrent event occurring

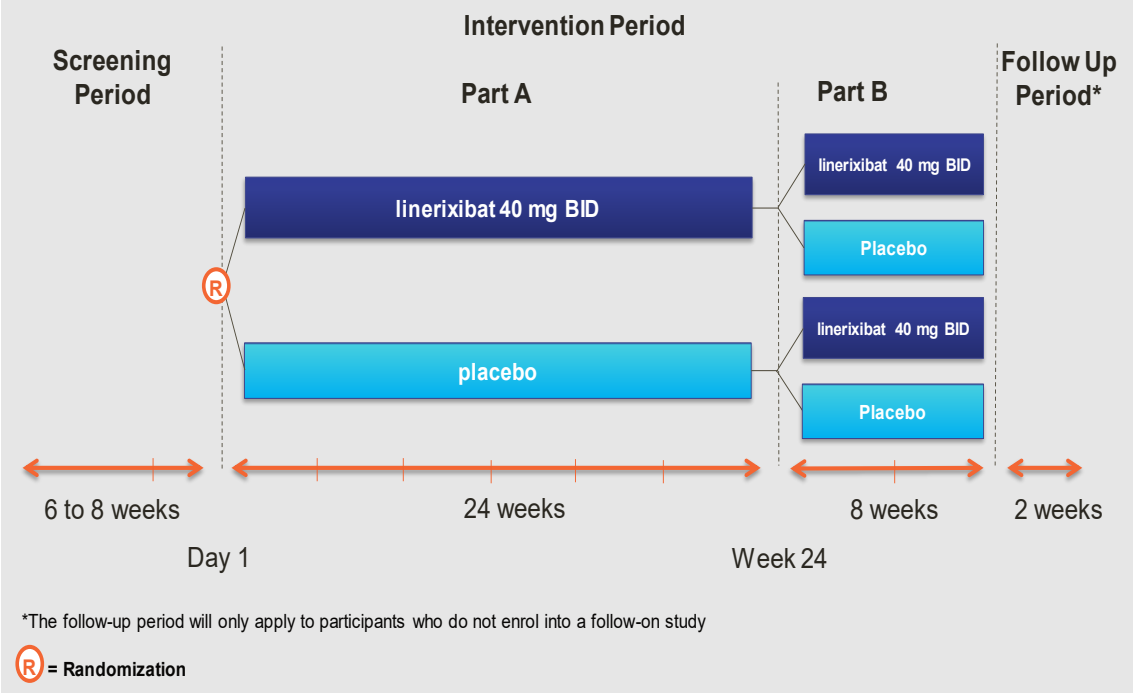
	Safety Endpoints Including: • **Adverse Events (AEs) and Serious Adverse Events (SAEs)**	**Safety Endpoints Including:** • **Vital signs** • **12-lead ECG** • **Clinical laboratory evaluation (including liver chemistry panel and fasting lipids)**
COVID-19 pandemic)		
Change in background itch therapy or use of rescue medication	Addressed with treatment policy strategy, i.e., regardless of the intercurrent event occurring	Addressed with treatment policy strategy, i.e., regardless of the intercurrent event occurring

Supplementary Estimand for the Safety Objective

A supplementary estimand for the safety endpoints Adverse Events (AEs) and Serious Adverse Events (SAEs) will be defined to assess treatment discontinuation for any reason with the treatment policy strategy, i.e., regardless of the intercurrent event occurring.

1.2. Study Design

Overview of Study Design and Key Features



Design Features	This study is a Phase 3, multicenter, 2-part (Part A and Part B), randomized, double-blind, placebo-controlled study in PBC patients with moderate to severe pruritus. Part A will evaluate the efficacy,

Overview of Study Design and Key Features	
	safety and impact on health-related quality of life of linerixibat compared with placebo over 24 weeks, while Part B will assess the return of itch over 8 weeks after withdrawal of linerixibat. In both Part A and Part B, participants may be treatment naïve, have had prior itch therapy and/or may take stable background itch therapy at the discretion of the physician in the best interest of the patient, or rescue medication where applicable. This study consists of four Periods: Screening, Intervention (Part A and Part B) and Follow-up. Screening Period: Participants will attend two screening visits to assess eligibility as described below: Screening 1 (Visit 1): Assessments to be performed as described in the schedule of assessments (SoA). Itch severity will be assessed retrospectively based on participant recall and then prospectively for eligibility purposes using a twice daily eDiary throughout the screening period. Screening 2 (Visit 2): Additional liver biochemistry assessments to be performed at least 4 weeks after screening visit 1 to ensure stability of liver disease for participant safety and eligibility purposes. Intervention (Treatment) Period: will start with Part A (Day 1 through Week 24) followed by Part B (Week 24 through Week 32). Eligible participants will be randomized in a 1:1:1:1 ratio to receive linerixibat 40 mg BID in Part A and Part B, linerixibat 40 mg BID in Part A and placebo in Part B, placebo in Part A and Part B, or placebo in Part A and linerixibat 40 mg BID in Part B. Follow-up Period or Follow-on Study: Participants who complete treatment in Part A and Part B will be offered the opportunity to take part in a separate long-term follow-on study, where linerixibat will be provided in an open-label manner. This open-label, long-term study will assess safety and tolerability (and efficacy in participants transferring from this study). Participants who do not enter the follow-on study will have a follow-up phone call approximately 7-14 days after the last dose of study drug. This study will be delivered through a partially decentralized clinical trial approach. Decentralized clinical trials (DCTs) are defined as trials executed through telemedicine, mobile/local healthcare providers and/or mobile technologies [Apostolaros, 2020]. This study will employ a combination of two types of sites: 1) Metasite, and 2) brick and mortar sites, which may employ decentralized methodologies. The sites and operating models are defined as follows:

Overview of Study Design and Key Features	
	Metasite: Participants at the Metasite will complete all visits remotely (Metasite model). The Metasite model uses telemedicine-based clinical research to allow all trial visits to be conducted remotely. Brick and mortar sites: Brick and mortar sites may offer participants the opportunity to complete all visits at the site (brick and mortar model) or to complete some visits remotely (flexible model). Under the flexible model, consenting/initial screening and dispensing visits will be performed in clinic. All other visits may be performed remotely as described in the Study Reference Manual (SRM). The flexible model will be available worldwide in sites/countries where the regulations and infrastructure allow. This study plans to randomize approximately 230 participants with PBC and moderate to severe cholestatic pruritus. Participants receiving concomitant bile acid binding resins may comprise up to 15% of the overall Phase 3 study population.
Study intervention	After a Screening period, eligible participants will be enrolled in the study and start the 24-week Treatment Period Part A followed by 8 weeks of Treatment Period Part B. Finally, participants will be offered enrolment in a long-term follow-up study after Treatment Period Parts A and B. Participants who do not enter the follow-up study will have a 2-week follow-up. The total expected duration of study participation for each participant is approximately 40-42 weeks. Permanent Discontinuation: If study intervention is permanently discontinued following randomization and before the scheduled follow-up, participants should not be withdrawn from the study and should complete the remaining scheduled visits and all required assessments, including the eDiary. If the participant is withdrawing from the study, they will be encouraged to complete Early Discontinuation Visit (please refer to Schedule of Activities (SoA) in the study protocol) before their last dose of study intervention if possible. Restart of Study Intervention: Restart refers to resuming study intervention following liver stopping events in which there is a clear underlying cause (other than drug induced liver injury [DILI]) of the liver event (e.g. biliary obstruction, pancreatic events, hypotension, acute viral hepatitis). Furthermore, restart is not permitted following liver stopping event when the underlying cause was alcohol-related hepatitis. Concomitant Therapy: Concomitant therapy for PBC or any medications that may affect itch as shown in the study protocol are permitted but must be stable prior to Screening and throughout the Screening and Intervention Periods. Rescue medication is permitted

Overview of Study Design and Key Features	
	only for participants who are not receiving background cholestatic pruritus treatment when entering the study. If a participant reports experiencing worsening itch and requests rescue therapy, it may be considered only after 12 weeks of randomized treatment. Once the criteria for using rescue medication have been satisfied, the investigator may offer treatment with cholestyramine to the participant per treatment guidelines.

Study intervention Assignment	Intervention (treatment) period will start with Part A (Day 1 through Week 24) followed by Part B (Week 24 through Week 32). Day 1 is the day on which the first study intervention administration takes place and may differ from the randomization date. Eligible participants will be randomized in a 1:1:1:1 ratio to receive linerixibat 40 mg BID in Part A and Part B, linerixibat 40 mg BID in Part A and placebo in Part B, placebo in Part A and Part B, or placebo in Part A and linerixibat 40 mg BID in Part B. Participants who had initially been taking linerixibat in Part A will either continue on linerixibat or switch to placebo at Week 24 for Part B. Participants who had been taking placebo in Part A will either continue on placebo or switch to linerixibat at Week 24 for Part B.
	
	Randomization will be stratified by:
	
	• Severity of pruritus based on Monthly Itch Score for the 28 days preceding randomization:
	  • Moderate: ≥ 4 and < 7
	  • Severe: ≥ 7
	• Concomitant cholestatic pruritus treatment regimen (definitions of cholestatic pruritus treatments can be found in the main protocol text):
	  • Bile acid binding resin (BABR)-containing regimen
	  • Regimen that does not contain BABRs
	  • No defined cholestatic pruritus treatment
	• Region/site:
	  • US Brick and Mortar
	  • US Metasite
	  • Rest of world (RoW) Brick and Mortar
	
	Region/site will serve as an administrative stratification factor and will not be used in the analysis.

Interim Analyses	The study has a planned futility analysis to be conducted after approximately 100 participants reach 24 weeks of treatment (or discontinue early). The analysis will be overseen by an Independent

Overview of Study Design and Key Features	
	Data Monitoring Committee (IDMC) with non-binding futility guidelines as documented in the IDMC Charter (current version: GlaxoSmithKline Document Number TMF-16125967). The study may be terminated following this analysis if the probability of demonstrating clinical efficacy based on observed data is too low to support continuing or following the safety review if the overall risk-benefit is assessed to be unfavorable. In the event of the study being discontinued, GSK processes will be followed to ensure the best interests and the safety of the participants. The trial will not be stopped for early benefit at this point and no alpha-spend is incurred.
Primary and Final Analyses	The primary statistical analysis to assess the primary, secondary, and safety objectives will be performed when all randomized participants have completed the Part A intervention period (up to Week 24/Visit 9). In addition, available data for exploratory endpoints in Part A and Part B will be analyzed as part of this primary statistical analysis. By the time the last participant completes the Part A intervention period, it is anticipated that the majority of participants will have also completed the Part B intervention period (Week 24 through Week 32). Therefore, all available Part B data will also be assessed and reported as part of the primary statistical analysis. A final analysis will then be performed when all participants have completed the Part B intervention period, and the follow-up period, if applicable. This analysis will include complete Part B data and the full data for exploratory endpoints in Part A for which full datasets were not available at the time of the primary analysis.

2. STATISTICAL HYPOTHESES

2.1. Primary Endpoint

The primary objective is to investigate the superiority of 24 weeks of treatment with oral linerixibat compared with placebo on itch in PBC patients with cholestatic pruritus. The primary endpoint is Change from Baseline in MIS over 24 weeks using a 0-10 NRS. The primary estimand is the difference in means between linerixibat and placebo in the change from baseline in itch score over 24 weeks in PBC patients with cholestatic pruritus. The primary analysis will be a mixed model approach based on MIS over 24 weeks (Weeks 4, 8, 12, 16, 20, and 24). The primary hypothesis is that linerixibat reduces itch score over 24 weeks compared to placebo. A negative change from baseline (or reduction) represents an improvement in itch.

The null hypothesis (H_0) assumes that there is no difference in the reduction of itch score over 24 weeks between linerixibat and placebo groups. The alternate hypothesis (H_1) assumes that there is a difference in the reduction of itch score over 24 weeks between linerixibat and placebo groups.

The overall significance level for the primary endpoint analysis is set at the two-sided 0.05 level.

2.2. Secondary Endpoints and Multiplicity Control

Formal multiplicity-controlled testing will be conducted for the key secondary endpoints of change from baseline in WIS at Week 2, change from baseline in MSS over 24 weeks, as well as responder endpoints defined as achieving ≥2, ≥3, and ≥4-point reduction from Baseline in the MIS at Week 24. Additional non-controlled testing will be conducted for supportive secondary endpoints including change from baseline in PBC-40 domain scores at Week 24, change from baseline in PGI-S over 24 weeks, PGI-C over 24 weeks, change from baseline in ALP at Week 24, and change from baseline in bilirubin at Week 24.

A step-down/hierarchical approach will be used to test the following key secondary endpoints if statistical significance is achieved for the primary endpoint:

- Change from Baseline in WIS at Week 2
- Change from Baseline in MSS over 24 weeks
- Responders defined as achieving a ≥2-point reduction from Baseline in the MIS at Week 24.
- Responders defined as achieving a ≥3 -point reduction from Baseline in the MIS at Week 24.
- Responders defined as achieving a ≥4 -point reduction from Baseline in the MIS at Week 24.

These tests will be conducted sequentially at a 5% two-sided significance level. Other secondary endpoints will also be tested but outside this formal hierarchy.

3. ANALYSIS SETS

Analysis Set	Definition / Criteria	Analyses Evaluated
Screening	• All participants who sign the Informed Content Form (ICF)	• Study Population
Enrolled	• All participants who sign the ICF and pass screening (meet the eligibility criteria)	• Study Population
Intent to Treat (ITT)	• This population will comprise all randomized participants. Participants in the ITT Population will be classified according to the treatment as randomized.	• Study Population • Efficacy • Biomarker

Analysis Set	Definition / Criteria	Analyses Evaluated
	• For Part B, this will be all participants in the ITT who do not withdraw in Part A.	
Non-BABR ITT	• The Non-BABR ITT Population is a subset of the ITT population who are not receiving concomitant bile acid binding resin therapy (cholestyramine, colesevalem, colestimide, and/or colestipol) at randomization.	• Efficacy
Non-Metasite ITT	• The non-Metasite ITT population is a subset of the ITT population who are not enrolled at the Metasite.	• Efficacy
Ursodeoxycholic acid (UDCA) ITT	• A subset of the ITT population who are receiving UDCA at randomization	• ALP
ALP ≥1.67x ULN ITT	• A subset of the ITT population who have serum alkaline phosphatase [ALP] concentrations ≥1.67x upper limit of normal [ULN] at baseline.	• ALP • Total Bilirubin
Safety	• All randomized participants who take at least 1 dose of study intervention. Participants will be analyzed according to the treatment they actually received.* • For Part B, this will be all participants in the safety population who take at least 1 dose in Part B.	• Safety • Study Population
PK	• Any randomized participant who had at least 1 PK sample. Participants will be analyzed according to the treatment they actually received.	• PK

* In cases where incorrect treatment was received for the duration of a treatment period (i.e., entire Part A or entire Part B), then the participant will be analyzed according to the treatment received. If the participant received incorrect treatment for only part of a treatment period and therefore received both treatments during the same treatment period, then analysis will be based on active treatment.

4. STATISTICAL ANALYSES

4.1. General Considerations

4.1.1. General Methodology

Planned treatment will be used for the analyses except safety and PK, which will use actual treatment. Participants who prematurely withdrew from study will not be replaced. In the case of wrong stratification assigned at the time of randomization, the analyses will be performed based on the actual stratum per data collected in the CRF.

Confidence intervals (CIs) will use 95% CIs unless otherwise specified. Where requested, *t*-type confidence limits will be constructed for LS means estimates, unless otherwise specified. The alpha level for all the tests is at two sided 0.05 unless otherwise specified.

Unless otherwise specified, continuous data will be summarized using descriptive statistics: n, mean, standard deviation (SD), median, minimum and maximum. Categorical data will be summarized as the number and percentage of participants in each category.

For descriptive analyses, intercurrent event strategy will not be applied. In other words, data impacted by an intercurrent event will not be discarded for descriptive summary displays that have no statistical model involved.

For descriptive summary of responder rate, the responder at Week 4, 8, 12, 16, 20, 24 will be dichotomized from the corresponding MIS. Participants with missing responder data at the Week of interest due to the missing of MIS for the 4 weeks prior will be classified as non-responders for that Week.

It is anticipated that participant accrual will be spread thinly across centers; therefore, summaries of data by center are unlikely to be informative and will not be provided.

For all the analyses using MMRM, an unstructured covariance structure will be assumed; if this does not converge then compound symmetry (CS) will be used, followed by an autoregressive [AR(1)] covariance structure if CS does not converge. The degrees of freedom will be computed using the Kenward-Roger (KR) method. Distributional assumptions underlying the MMRM will be examined by obtaining a normal probability plot of the residuals and a plot of the residuals versus the fitted values (i.e. checking the normality assumption and constant variance assumption of the model, respectively) to gain confidence that the model assumptions are reasonable. If there are any clear departures from the distributional assumptions, alternative models will be explored using appropriate transformed data. When the least-squares (LS) means over multiple timepoints are estimated for each treatment group and the difference from placebo is derived, equal weight is applied to each timepoint and simple average of LS means at each time point is used.

4.1.2. Baseline Definition

The summary of baseline definitions for all the study endpoints, except aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), total bilirubin (TB), Total serum bile acids, Serum C4, Serum autotaxin, and FGF-19 is shown in Table 1.

Baseline serum AST, ALT, ALP and TB values will be established using an average of two sets of laboratory values as detailed in protocol Section 8.3.4.1. The two samples are expected to be obtained at least 4 weeks apart; however, in cases where this was not feasible, both samples will still be used for the calculation of the baseline value. The baseline value will be established using the average of Samples 1 and 2. In cases where at least one of the two samples is abnormal and the variance between Sample 1 and Sample 2 was > 40% and a third sample was obtained confirming eligibility prior to randomization, then the average of Samples 1 and 3 will be used, as described in Section 8.3.4.1 of the protocol. In cases where only one value is present, then this value will be used as baseline. If any values are below the LOQ and are therefore reported as a non-numeric value (e.g., < 3), the rules in Section 6.3.9 will be applied.

For PD/Biomarker samples (Total serum bile acids, Serum C4, Serum autotaxin, FGF-19, and IL-31) baseline value will be established using the average of the Screening 2 and last assessment prior to the first dose values. In cases where only one of these values is present then this value will be used as baseline. For analysis of PD/Biomarker samples in Part B, the same baseline as Part A will be used.

For other laboratory results, vital signs, GSRS, PBC-40, PGI-S, ESS, BDI-II, ELF, Fibroscan, and ECG endpoints in Part A, the baseline assessment will be the last assessment conducted prior to the first dose of randomized medication. For the analysis of PBC-40, PGI-S, ESS, BDI-II, ELF during Part B, the baseline will be the assessment performed at Visit 9.

For safety related endpoints in Part B, including vital signs, ECG, clinical laboratory evaluations, and GSRS, the same baseline for Part A will be used.

For daily data collected on the eDiary, the baseline will be as follows:

- **WIS** at baseline will be the average of the WDI scores i.e. the higher of the AM and PM score each day, in the 7 days prior to randomization
- **MIS at baseline for Part A** will be the worst WIS in the 28 days immediately preceding randomization. More specifically, for each week of the 4 weeks immediately preceding randomization, the WIS will be calculated, and the worst WIS of the 4 weeks will be the baseline.
- **MIS at baseline for Part B** will be the worst WIS of the 28 days immediately preceding Part B. More specifically, for each week of the 4 weeks immediately preceding Part B, the WIS will be calculated, and the worst WIS of the 4 weeks will be the baseline.

- **MSS at baseline for Part A** will be the worst WSS of the 28 days immediately preceding randomization. More specifically, participant's sleep quality is recorded on the eDiary each morning using a 0-10 NRS. For each week of the 4 weeks immediately preceding randomization, the WSS will be calculated, and the worst MSS of the 4 weeks will be the baseline.
- **MSS at baseline for Part B** will be the worst WSS of the 28 days immediately preceding Part B. More specifically, for each week of the 4 weeks immediately preceding Part B, the WSS will be calculated, and the worst WSS of the 4 weeks will be the baseline.
- **Monthly Fatigue Score (MFS) at baseline for Part A** will be the worst Weekly Fatigue Score (WFS) of 28 days immediately preceding randomization. More specifically, fatigue is recorded on the eDiary each evening using a 0-10 NRS. For each week of the 4 weeks immediately preceding randomization, the average of the Daily Fatigue Scores (DFS) in one week will be used to calculate WFS. The worst of WFS of the 4 weeks will be the baseline.
- **MFS at baseline for Part B** will be the worst WFS of the 28 days immediately preceding Part B. More specifically, for each week of the 4 weeks immediately preceding Part B, the WFS will be calculated, and the worst WFS of the 4 weeks will be the baseline.

Table 1 Baseline definition

Parameter	Study Assessments Considered as Baseline				
	Last assessment prior to the first dose of Part A[1]	Mean of 7 days prior to randomization	Monthly Score for 28 days prior to randomization	Monthly Score for 28 days prior to day 1 of Part B	Visit 9 (Week 24)
MIS, MSS, MFS at Part A			X		
MIS, MSS, MFS at Part B				X	
WIS		X			
PBC-40, PGI-S, ESS, BDI-II at Part A	X				
PBC-40, PGI-S, ESS, BDI-II at Part B					X
GSRS	X				
ECGs	X				

Parameter	Study Assessments Considered as Baseline				
	Last assessment prior to the first dose of Part A[1]	Mean of 7 days prior to randomization	Monthly Score for 28 days prior to randomization	Monthly Score for 28 days prior to day 1 of Part B	Visit 9 (Week 24)
Laboratory data excluding AST, ALT, ALP, and TB[2]	X				
Vital signs	X				
ELF (Part A) Fibroscan	X				
ELF at Part B					X

[1] If time is not collected, Day 1 assessments are assumed to be taken prior to first dose.
[2] Please see the baseline definition of AST, ALT, ALP, and TB in the text above.

4.2. Primary Endpoint Analyses

4.2.1. Definition of Endpoint/Estimands

The primary endpoint is Change from Baseline in MIS over 24 weeks using a 0-10 NRS. The MIS will be determined from the worst WIS for the month (4 weeks). A week will have a WIS if there are at least 4 WDI scores in the week, where a WDI score needs at least 1 entry for that day. Otherwise, the WIS for the week will be missing. If one or more WIS are missing, then the worst of the non-missing WIS is selected as the MIS. If no WIS are available for a month, the MIS will be missing.

The following intercurrent events have been considered for the evaluation of the primary endpoint of MIS:

- Permanent treatment discontinuation, disruptions in treatment, treatment delays unrelated to the COVID-19 pandemic, addressed with treatment policy strategy
- Permanent treatment discontinuation, disruptions in treatment, treatment delays related to the COVID-19 pandemic, addressed with hypothetical strategy
- Change in background itch therapy or use of rescue medication, addressed with treatment policy strategy

The number and percentage of participants experiencing each intercurrent event will be summarized.

The COVID-19 pandemic related intercurrent events are defined as permanent treatment discontinuation, disruptions in treatment, or treatment delays related to study treatment supply interruption due to COVID-19 pandemic, such as site closure, lockdowns, and other drug supply issues. The intercurrent events caused by COVID-19 infection or concern of COVID-19 are not included in the above definition of COVID-19 pandemic related and will be treated the same manner as the other COVID-19 pandemic unrelated intercurrent events.

Data following intercurrent events that are unrelated to COVID-19 will be included as available. WDI scores impacted by COVID-19 pandemic related intercurrent events will be discarded. For example, if treatment is disrupted during Day 28 to Day 30 due to COVID-19, the WDI scores collected during this period will be discarded to construct the corresponding WIS. If treatment is permanently discontinued from Day 30 due to COVID-19, the data from Day 30 onwards will be discarded, thus leading to missing WIS and consequently missing MIS.

The summary measure will be mean difference between treatment groups.

4.2.2. Main Analytical Approach

The primary analysis will be a mixed model for repeated measures (MMRM) of change from baseline in MIS over the 24-week treatment period. The MMRM will be fitted after the outcomes impacted by COVID-19 pandemic related intercurrent events have been discarded. This assumes that unobserved outcomes of participants who withdrew from the study or who experienced a COVID-19 pandemic related intercurrent event are missing at random (MAR).

Endpoint / Variables
• Change from Baseline in Monthly Itch Scores (MIS) over 24 weeks using a 0-10 numerical rating scale (NRS)
Model Specification
• Mixed Models Repeated Measures (MMRM) analysis. • Terms in the model: $Y = Treatment + Visit + Treatment * Visit + BL_{Itch} + BL_{Itch} * Visit + BL_{Concomitant\ itch\ med} + Error$ ○ **Response** (Y): change from baseline in MIS at Visits 4 – 9 of the study intervention during Part A. ○ **Categorical**: treatment group, visit, baseline concomitant itch therapy (3-level; Regimen contains BABR, Regimen that does not contain BABR, No defined anti-cholestatic pruritus treatment) ○ **Continuous**: baseline MIS ○ **Interaction**: treatment group*visit, baseline MIS*visit ○ **Repeated**: visit Visit refers to the planned visits (V4 – V9) at week 4, 8, 12, 16, 20, and 24 of the study intervention during Part A. See Section 6.3.7.1. • Baseline is defined in Section 4.1.2
Model Results Presentation
• The LS means and the corresponding 95% confidence intervals of change from baseline in MIS over 24 weeks (equal weight for each time point, i.e., average of Weeks 4, 8, 12, 16, 20 and 24 LS means) for

linerixibat and placebo will be reported. The LS mean, 95% CI, and p-value for the difference from placebo will be reported in the same table. • The LS means and 95% confidence intervals for change from baseline in MIS at each visit (Weeks 4, 8, 12, 16, 20, and 24) will be reported for linerixibat and placebo, as well as the LS means and corresponding 95% CIs for the difference from placebo at each visit.
Model Checking & Diagnostics
• See Section 4.1.1
Subgroup/covariate Analyses
○ The MMRM analysis will be repeated by including each of the covariates indicated in Section 4.6.1, respectively in separate models, along with its interaction with treatment group. For continuous covariates, the model estimates are obtained for the median, lower, and upper quartiles of the covariate. For covariates concomitant background itch therapy (Yes vs. No), baseline concomitant fibrates (Yes vs. No), baseline concomitant BABR and baseline cholestyramine use, baseline concomitant itch therapy (3-level) will not be included in the model (i.e., we will remove the baseline concomitant itch therapy [3-level] variable in place of these variables in each respective model to avoid confounding). ○ The LS mean of change from baseline and its corresponding 95% CI will be reported for each treatment group, as well as the LS mean and its 95% CI of difference from placebo. The p-values for treatment group, covariate, and interaction between treatment group and covariate will be reported. ○ Additionally, a shrinkage estimation may be applied to the estimate and confidence intervals for the subgroups, further details are provided in Section 4.6.4.
Supplementary Analyses
• The same MMRM analysis will be applied after discarding data impacted by the intercurrent events. The LS means and the corresponding 95% confidence intervals of change from baseline in MIS over 24 weeks for linerixibat and placebo will be reported. The LS mean, 95% CI, and p-value for the difference from placebo will be reported in the same table. See Section 4.2.4 for more details. • The same MMRM analysis will be applied to the non-bile acid binding resin ITT Population. The LS means and the corresponding 95% confidence intervals of change from baseline in MIS over 24 weeks for linerixibat and placebo will be reported. The LS mean, 95% CI, and p-value for the difference from placebo will be reported in the same table. See Section 4.2.4 for more details. • The same MMRM analysis will also be applied to the non-Metasite ITT Population. The LS means and the corresponding 95% confidence intervals of change from baseline in MIS over 24 weeks for linerixibat and placebo will be reported. The LS mean, 95% CI, and p-value for the difference from placebo will be reported in the same table. See Section 4.2.4 for more details. • The impact of operating model will be tested by including operating model and the interaction between operating model and treatment group into the MMRM. See Section 4.6.3 for the details of operating model definition, analysis model, and displays. • The empirical cumulative distribution function (eCDF) graph by treatment group of change from baseline in non-missing MIS averaged over Week 4, 8, 12, 16, 20, and 24 (equal weight for the timepoints) will be provided. The observed number of observations will be used as the denominator. • Additionally, the empirical cumulative distribution function (eCDF) graph by treatment group of change from baseline in Imputed MIS averaged over Week 4, 8, 12, 16, 20, and 24 (equal weight for the timepoints) will be provided. Multiple imputation will be applied to impute the missing MIS by treatment groups. Please see step 1 in Section 4.2.3 for details of the imputation model.

4.2.3. Sensitivity Analyses

A sensitivity analysis will be conducted to examine the sensitivity of inferences to departure from the MAR assumption if the primary endpoint is significant in favor of linerixibat. In this sensitivity analysis, a tipping point approach will be used. This is akin to progressive stress-testing to assess how severe departures from the MAR assumption must be in order to change the overall conclusion of the primary analysis. The following steps will be taken:

1. Missing MIS for Week 4, 8, 12, 16, 20, and 24, including data discarded due to COVID 19 related intercurrent events, will be imputed under the MAR assumption.
   i. Briefly, the first step of imputation will use SAS MI procedure to impute the intermitted missing values and the variables to be used are MIS at baseline, Week 4, 8, 12, 16, 20, and 24, and treatment group. Treatment group will be included in the SAS by statement in the first instance (i.e., imputed separately for each treatment group), in case of non-convergence then it may be included as a covariate (using the SAS PROC MI var statement). The seed number for the imputation is 212620. The number of imputations is 100. Since the MIS NRS has a 0-10 range, any imputed values that are less than 0 will be set to 0, and any values greater than 10 will be set to 10.
   ii. The 2nd step of imputation will use the regression method to impute the monotone missing values, for each of the monotone missing dataset (out of the 100 imputed as indicated above). The variables to be used are baseline concomitant itch therapy, as well as MIS at baseline, Week 4, 8, 12, 16, 20, and 24, and treatment group. Treatment group will be included in the SAS by statement in the first instance (i.e., imputed separately for each treatment group), and a categorical covariate in the case of non-convergence (using the SAS PROC MI class statement). The seed number for the imputation is 212620. Since the MIS NRS has a 0-10 range, any imputed values that are less than 0 will be set to 0, and any values greater than 10 will be set to 10. Then, change from baseline MIS will be calculated.
2. The 100 complete data sets generated from step 1 will be analyzed by the MMRM model specified in Section 4.2.2, respectively.
3. The results from the MMRM analyses of 100 complete data sets, including the LS mean and its standard error of difference from placebo in change from baseline in MIS over 24 weeks, will be combined using SAS PROC MIANALYZE based on Rubin's rules to have a pooled estimate of LS mean and 95% CI of difference from placebo, as well as the p-value.
4. Repeat step 1, with a pair of shift parameters applied to the imputed datapoints for linerixibat and placebo groups, respectively. As described above, any values that are less than 0 will be set to 0, and any values greater than 10 will be set to 10.
5. Repeat steps 2 – 3 to obtain the combined estimate of difference from placebo with 95% CI and associated p-value.
6. Repeat steps 4 – 5 with another pair of shift parameters.

Overall, a sequence of shift parameters ranging from -3 to 3 by 0.5 increments will be applied to find the tipping point where the result is no longer statistically significant. The shift parameters will vary independently for each arm and will include scenarios where participants in the linerixibat arm have worse outcomes following withdrawal from the study than participants in the placebo arm. The shift parameters will only be applied to the imputed monotone missing values, but not to the intermittent missing values. The tipping point results will be reported in a 2-dimensional table that includes the combined estimate of difference from placebo in change from baseline in MIS over 24 weeks, as well as its 95% CI and corresponding p-value, for each pair of shift parameters.

4.2.4. Additional Estimands

As a supplementary estimand for the primary objective, the hypothetical treatment effect of linerixibat will be compared to placebo in the absence of any intercurrent events. Hypothetical strategy will be applied to all the intercurrent events defined in Section 4.2.1, including treatment discontinuation, disruptions in treatment or treatment delays, changes in itch therapy or use of rescue medication, regardless of whether an intercurrent event is related to COVID-19 or not. The WDIs impacted by any intercurrent events will be discarded in the same way as detailed in Section 4.2.1, and the MMRM will then be fitted. This assumes that unobserved outcomes of participants who withdrew from the study or who experienced an intercurrent event are MAR.

The main analytical approach will be repeated using the non-bile acid binding resin ITT population. Any participant taking bile acid binding resin as the background itch therapy at baseline will be excluded from the non-bile acid binding resin ITT population and any data from these excluded participants will not contribute to the analysis.

The main analytical approach will also be repeated by using the non-Metasite ITT population. Any participants enrolled at Metasite will be excluded from the non-Metasite ITT population and any data from these excluded participants will not contribute to the analysis.

4.3. Secondary Endpoints Analyses

4.3.1. Key/Confirmatory Secondary Endpoints

4.3.1.1. Definition of Endpoints/Estimands

- Change from baseline in Weekly Itch score at Week 2

The WIS defined earlier at Week 2 will be evaluated. Similar to the primary endpoint, treatment policy strategy will be used for the intercurrent events unrelated to the COVID-19 pandemic, whereas hypothetical strategy will be used for the intercurrent events related to the COVID-19 pandemic. The discard of data impacted by COVID-19 intercurrent events is detailed in Section 4.2.1.

- Change from Baseline in Monthly Sleep Score over 24 weeks:

The MSS, based on a 0-10 NRS, will be the worst WSS available for the month (4 weeks). A week will have a WSS if at least 4 days in the week have a score. Otherwise, WSS will be missing. If one or more WSS are missing, then the worst of the non-missing WSS is selected as the worst for that month. If no WSS are available for a month, the MSS will be missing.

Treatment policy strategy will be used for the intercurrent events unrelated to the COVID-19 pandemic, whereas hypothetical strategy will be used for the intercurrent events related to the COVID-19 pandemic. The DSS affected by the COVID-19 pandemic related intercurrent events will be discarded to construct the corresponding WSS, which may lead to missing WSS and consequently missing MSS.

- Responder defined as achieving a ≥2-point reduction from baseline in the Monthly Itch score at Week 24;
- Responder defined as achieving a ≥3-point reduction from baseline in the Monthly Itch score at Week 24;
- Responder defined as achieving a ≥4-point reduction from baseline in the Monthly Itch score at Week 24:

Responder will be evaluated using three different thresholds based on change from baseline in MIS at Week 24, which is the worst WIS of weeks 21 - 24. Treatment policy strategy will be used for the intercurrent events unrelated to the COVID-19 pandemic, whereas hypothetical strategy will be used for the intercurrent events related to the COVID-19 pandemic. WDI scores affected by COVID-19 pandemic related intercurrent events will be discarded, which may lead to missing WIS and consequently missing MIS. If there are missing MIS, multiple imputation will be applied to impute MIS using available data for all participants. The imputed MIS at Week 24 and the corresponding baseline score will then be dichotomized to form the responder endpoint.

4.3.1.2. Main Analytical Approach of Secondary Endpoints

Change from baseline WIS at Week 1 and 2 will be fitted to a MMRM model including baseline WIS as a continuous covariate, treatment, week (Weeks 1 and 2), concomitant use of itch medications at baseline, treatment by week interactions, and baseline WIS by week interaction as fixed effects. The comparison at Week 2 will be estimated from the MMRM model.

Endpoint / Variables
• Change from baseline in Weekly Itch score at Week 2
Model Specification
• Mixed Models Repeated Measures (MMRM) analysis. • Terms in the model: $Y = Treatment + Week + Treatment * Week + BL_{Itch} + BL_{Itch} * Week + BL_{Concomitant\ itch\ med} + Error$ ○ **Response** (Y): change from baseline in WIS at Week 1 and 2 of the study intervention during part A. ○ **Categorical**: treatment group, week, baseline concomitant itch therapy (3-level) ○ **Continuous**: baseline WIS ○ **Interaction**: treatment group*week, baseline WIS*week ○ **Repeated**: week In this model Week is used rather than visit since Week 2 does not correspond to a visit. • Baseline is defined in Section 4.1.2.
Model Checking & Diagnostics
• See Section 4.1.1
Model Results Presentation
• The LS means and the corresponding 95% confidence intervals of change from baseline in WIS at Week 1 and 2 for linerixibat and placebo will be reported. The LS mean, 95% CI, and p-value for the difference from placebo will be reported in the same table.

Supplementary Analyses
• The same MMRM analysis will be applied to the Non-bile acid binding resin ITT Population. The same model results display as the ITT population will be produced.

Change from baseline of MSS at planned Visits 4 - 9 will be analyzed using a MMRM model including baseline MSS as covariate, treatment, visit, concomitant use of itch medications at baseline, treatment by visit interaction, and baseline MSS by visit interaction as fixed effects. The comparison of effects averaged over the visits will be estimated from the MMRM.

Endpoint / Variables
• Change from baseline in Monthly Sleep score over 24 weeks
Model Specification
• Mixed Models Repeated Measures (MMRM) analysis. • Terms in the model: $Y = Treatment + Visit + Treatment * Visit + BL_{Sleep} + BL_{Sleep} * Visit + BL_{Concomitant\ itch\ med} + Error$ ○ **Response** (Y): change from baseline in MSS at Visits 4 – 9 of the study intervention during part A. ○ **Categorical**: treatment group, visit, baseline concomitant itch therapy (3-level) ○ **Continuous**: baseline MSS ○ **Interaction**: treatment group*visit, baseline MSS*visit ○ **Repeated**: visit Visit refers to the planned visits (V4 – V9) at week 4, 8, 12, 16, 20, and 24 of the study intervention during part A. • Baseline is defined in Section 4.1.2.
Model Checking & Diagnostics
• See Section 4.1.1
Model Results Presentation
• The LS means and the corresponding 95% confidence intervals of change from baseline in MSS over 24 weeks for linerixibat and placebo will be reported. The LS mean, 95% CI, and p-value for the difference from placebo will be reported in the same table. • The LS means and 95% confidence intervals for change from baseline in MSS at each visit (Weeks 4, 8, 12, 16, 20, and 24) will be reported for linerixibat and placebo, as well as the LS means and corresponding 95% CIs for the difference from placebo.
Supplementary Analyses
• The same MMRM analysis will be applied to the Non-bile acid binding resin ITT Population. The same model results display as the ITT population will be produced. • A figure displaying Bland-Altman repeated measures correlations of change from baseline in MSS and MIS across all treatment groups will be produced, including all visits in Part A. The correlation coefficient will be displayed within the figure.

The summary measure for the responder endpoint will be the difference in responder rates (proportion of responders) between linerixibat and placebo groups, analyzed using a stratified analysis with Cochran-Mantel-Haenszel (CMH) weights, adjusting for baseline itch severity and concomitant use of itch medications.

Endpoint / Variables
• Responder defined as achieving a ≥2-point reduction from baseline in the Monthly Itch score at Week 24; • Responder defined as achieving a ≥3-point reduction from baseline in the Monthly Itch score at Week 24; • Responder defined as achieving a ≥4-point reduction from baseline in the Monthly Itch score at Week 24
Model Specification
• Each responder endpoint will be analyzed using a stratified analysis with Cochran-Mantel-Haenszel (CMH) weights, adjusting for itch severity at baseline (moderate: ≥ 4 and < 7; severe: ≥ 7) and concomitant therapy use at baseline (BABR-containing regimen, regimen that does not contain BABR, No cholestatic pruritus treatment). • The CMH estimate of the adjusted treatment difference will be calculated as a weighted average of strata-specific estimates of the treatment difference calculated within each of the following six Baseline analysis strata: o Moderate itch severity AND BABR-containing regimen o Moderate itch severity AND non-BABR-containing regimen o Moderate itch severity AND no cholestatic pruritus treatment o Severe itch severity AND BABR-containing regimen o Severe itch severity AND non-BABR-containing regimen o Severe itch severity AND no cholestatic pruritus treatment • The CMH estimate is given by $$\hat{d}_{cmh} = \frac{\sum W_k \hat{d}_k}{\sum W_k}$$ where y_k and x_k are the number of responder in linerixibat and placebo arms, respectively, in the *k*th strata; n_k is the number of participants in the linerixibat arm, m_k is the number of participants in the placebo arm, and $N_k = n_k + m_k$ is the total number of participants in the *k*th stratum; $$W_k = \frac{n_k m_k}{N_k}$$ are the CMH weight for the *k*th strata; and $$\hat{d}_k = \frac{y_k}{n_k} - \frac{x_k}{m_k}$$ are the estimates of the differences in response proportions between the two treatment arms, for the *k*th strata. • The corresponding two-sided 95% CI will be calculated as $$\hat{d}_{cmh} \pm 1.96 \times \sqrt{\hat{\text{var}}(\hat{d}_{cmh})}$$ where the variance estimator is consistent in both sparse data and large strata and is given below $$\hat{\text{var}}\left(\hat{d}_{cmh}\right) = \frac{\hat{d}_{cmh}\left(\sum P_k\right) + \sum Q_k}{\left(\sum n_k m_k / N_k\right)^2} = \frac{\hat{d}_{cmh}\left(\sum P_k\right) + \sum Q_k}{\left(\sum W_k\right)^2}$$ where

$$P_k = \frac{n_k^2 y_k - m_k^2 x_k + n_k m_k (m_k - n_k)/2}{N_k^2}$$

$$Q_k = \frac{x_k (m_k - y_k)/N_k + y_k (n_k - x_k)/N_k}{2}$$

Confidence intervals are readily available in SAS software (SAS Help Center: Common Risk Difference)

Multiple Imputation

- Multiple imputation will be applied to impute the missing MIS by treatment groups before the dichotomization of change from baseline in MIS at Week 24 to responder. Please see step 1 in Section 4.2.3 for details. The 100 complete data sets will then be analyzed by CMH method detailed above, and subsequently combined to estimate the CMH adjusted difference in responder rates between linerixibat and placebo groups.

Model Results Presentation

- The CMH estimate of the difference in the proportion of responders between the two treatment groups, corresponding 95% confidence interval, and p-value will be displayed for each responder definition.

Supplementary Analyses

- The same analysis will be applied to the Non-bile acid binding resin ITT Population. The same model results display as the ITT population will be produced.
- The empirical cumulative distribution function (eCDF) graph by treatment group of change from baseline in non-missing MIS at Week 24 will be provided. The observed number of observations will be used as the denominator.
- Additionally, the empirical cumulative distribution function (eCDF) graph by treatment group of change from baseline in Imputed MIS at Week 24 will be provided. Multiple imputation will be applied to impute the missing MIS by treatment groups. Please see step 1 in Section 4.2.3 for details of the imputation model.

4.3.2. Supportive Secondary Endpoint(s)

For supportive secondary endpoints, all the intercurrent events will be treated with treatment policy strategy.

- Change from Baseline in PBC-40 domain scores at Week 24

Change from baseline in PBC-40 domain scores at Visits 4 – 9 in Part A will be analyzed using a MMRM. The comparison at week 24 (V9) by domain will then be estimated from the MMRM model. Please refer to Section 1.1 for estimand strategies.

Endpoint / Variables

- Change from Baseline in PBC-40 domain scores at Week 24

Model Specification

- Mixed Models Repeated Measures (MMRM) analysis.
- Terms in the model:

$$Y = Treatment + Visit + Treatment * Visit + BL_{PBC40} + BL_{PBC40} * Visit + BL_{Concomitant\ itch\ med} + Error$$

  - **Response** (Y): change from baseline in PBC-40 domain scores at Visits 4 – 9 of the study intervention during Part A.
  - **Categorical**: treatment group, visit, baseline concomitant therapy (3-level)
  - **Continuous**: baseline PBC-40 domain score
  - **Interaction**: treatment group*visit, baseline PBC-40 domain scores*visit
  - **Repeated**: visit
- Baseline is defined in Section 4.1.2

Model Checking & Diagnostics
• See Section 4.1.1
Model Results Presentation
• The LS means and the corresponding 95% confidence intervals of change from baseline in PBC-40 at Week 24 (V9) for linerixibat and placebo, as well as the LS mean, 95% CI, and p-value for the difference from placebo will be reported.

- Change from baseline in PGI-S over 24 weeks

Change from baseline in derived PGI-S at visits 4 - 9 will be analyzed using MMRM. The comparison averaged over the 6 visits (24 weeks) will be estimated from the MMRM. Please refer to Section 1.1 for estimand strategies.

Endpoint / Variables
• Change from Baseline in PGI-S over 24 weeks
Model Specification
• Mixed Models Repeated Measures (MMRM) analysis. • Terms in the model: $Y = Treatment + Visit + Treatment * Visit + BL_{PGI-S} + BL_{PGI-S} * Visit + BL_{Concomitant\ itch\ med} + Error$ ○ **Response** (Y): change from baseline in PGI-S at Visits 4 – 9 of the study intervention during Part A. ○ **Categorical**: treatment group, visit, baseline concomitant itch therapy (3-level) ○ **Continuous**: baseline PGI-S ○ **Interaction**: treatment group*visit, baseline PGI-S*visit ○ **Repeated**: visit • Baseline is defined in Section 4.1.2
Model Checking & Diagnostics
• See Section 4.1.1
Model Results Presentation
• The LS means and the corresponding 95% confidence intervals of change from baseline in PGI-S over 24 weeks for linerixibat and placebo, as well as the LS mean, 95% CI, and p-value for the difference from placebo will be reported.

- PGI-C over 24 weeks

Derived PGI-C at Visits 4 – 9 will be analyzed using MMRM. The comparison averaged over the 6 visits (24 weeks) will be estimated from the MMRM. Please refer to Section 1.1 for estimand strategies.

Endpoint / Variables
• PGI-C over 24 weeks
Model Specification
• Mixed Models Repeated Measures (MMRM) analysis. • Terms in the model: $Y = Treatment + Visit + Treatment * Visit + BL_{Concomitant\ itch\ med} + Error$ ○ **Response** (Y): PGI-C at Visits 4 – 9 of the study intervention during Part A. ○ **Categorical**: treatment group, visit, baseline concomitant itch therapy (3-level) ○ **Interaction**: treatment group*visit ○ **Repeated**: visit

Model Checking & Diagnostics
• See Section 4.1.1
Model Results Presentation
• The LS means and the corresponding 95% confidence intervals of PGI-C over 24 weeks for linerixibat and placebo, as well as the LS mean, 95% CI, and p-value for the difference from placebo will be reported.
Supplementary analysis
• The percentage of participants with improved PGI-C will be reported by visit. Improved PGI-C will be defined as those who reported minimally improved, moderately improved, or very much improved. The number and percentage of participants will be reported for linerixibat and placebo, as well as the percentage difference, 95% CI, and p-value for the difference from placebo will be reported.

- Change from baseline in ALP at Week 24

Change from baseline in ALP at Visits 4 - 9 will be fitted by MMRM. The comparison at Visit 9 (week 24) will be estimated from the MMRM model. Please refer to Section 1.1 for estimand strategies.

Endpoint / Variables
• Change from Baseline in ALP at Week 24
Model Specification
• Mixed Models Repeated Measures (MMRM) analysis. • Terms in the model: $$Y = Treatment + Visit + Treatment * Visit + BL_{ALP} + BL_{ALP} * Visit + BL_{Concomitant\ itch\ med} + Error$$ ○ **Response** (Y): change from baseline in ALP at Visits 4 – 9 of the study intervention during part A. ○ **Categorical**: treatment group, visit, baseline concomitant itch therapy (3-level) ○ **Continuous**: baseline ALP ○ **Interaction**: treatment group*visit, baseline ALP*visit ○ **Repeated**: visit • Baseline is defined in Section 4.1.2
Model Checking & Diagnostics
• See Section 4.1.1
Model Results Presentation
• The LS means and the corresponding 95% confidence intervals of change from baseline in ALP at Week 24 for linerixibat and placebo, as well as the LS means, 95% CI, and p-value for the difference from placebo will be reported.
Supplementary analysis
• ALP and change from baseline ALP will be summarized descriptively by visit and treatment group for the participants with UDCA. • The same MMRM analysis will be applied to the participants in the ALP ≥1.67x ULN ITT Population. The LS means and the corresponding 95% confidence intervals of change from baseline in ALP at Week 24 for linerixibat and placebo will be reported. The LS mean, 95% CI, and p-value for the difference from placebo will be reported in the same table. In the event of small sample size and/or convergence issues, then only summary statistics will be reported.

- Change from baseline in bilirubin at Week 24

Change from baseline in bilirubin at Visits 4 – 9 will be fitted by MMRM. The comparison at visit 9 (week 24) will be estimated from the MMRM model. Please refer to Section 1.1 for estimand strategies.

Endpoint / Variables
• Change from Baseline in bilirubin at Week 24
Model Specification
• Mixed Models Repeated Measures (MMRM) analysis. • Terms in the model: $$Y = Treatment + Visit + Treatment * Visit + BL_{bilirubin} + BL_{bilirubin} * Visit + BL_{Concomitant\ itch\ med} + Error$$ ○ **Response** (Y): change from baseline in bilirubin at Visits 4 – 9 of the study intervention during Part A. ○ **Categorical**: treatment group, visit, baseline concomitant itch therapy (3-level) ○ **Continuous**: baseline bilirubin ○ **Interaction**: treatment group*visit, baseline bilirubin *visit ○ **Repeated**: visit • Baseline is defined in Section 4.1.2
Model Checking & Diagnostics
• See Section 4.1.1
Model Results Presentation
• The LS means and the corresponding 95% confidence intervals of change from baseline in bilirubin at Week 24 for linerixibat and placebo, as well as the LS mean, 95% CI, and p-value for the difference from placebo will be reported.
Supplementary analysis
• The same MMRM analysis will be applied to the participants in the ALP ≥1.67x ULN ITT Population. The LS means and the corresponding 95% confidence intervals of change from baseline in Bilirubin at Week 24 for linerixibat and placebo will be reported. The LS mean, 95% CI, and p-value for the difference from placebo will be reported in the same table. In the event of small sample size and/or convergence issues, then only summary statistics will be reported.

4.4. Exploratory Endpoints Analyses

4.4.1. Overview of Planned Analyses of Exploratory Endpoints

The hypothetical estimand strategy will not be applied for the analyses of exploratory endpoints. All the intercurrent events will be treated with treatment policy strategy. The table below provides an overview of the planned analyses. The ITT population will be used for the analysis unless otherwise specified. The details of statistical analyses are provided in Section 4.4.2.

Table 2 Overview of planned analyses of exploratory endpoints

Endpoints	Absolute							Change from Baseline						
	Stats Analysis			Summary		Individual		Stats Analysis			Summary		Individual	
	T	F	L	T	F	F	L	T	F	L	T	F	F	L
Gastrointestinal symptoms (Part A and Part B)														
Change from baseline in GSRS over time				Y							Y	Y		
Percentage of itch response days (Part A)														
≥2 point reduction				Y										
≥3 point reduction				Y										
≥4 point reduction				Y										

Endpoints	Absolute							Change from Baseline						
	Stats Analysis			Summary		Individual		Stats Analysis			Summary		Individual	
	T	F	L	T	F	F	L	T	F	L	T	F	F	L
Improvement, maintenance, or return of itch (Part B)														
Change from baseline (Part A) in WIS at 8 weeks				Y			Y				Y	Y		
Change from baseline (Part B) in MIS over 8 weeks				Y			Y				Y			
Improvement, maintenance or decline of quality of life (Part B)														
Change from baseline in PBC-40 domain scores at 8 weeks				Y			Y				Y			
Change from baseline in MSS over 8 weeks				Y							Y			
Change from baseline in PGI-S over 8 weeks				Y							Y			
PGI-C over 8 weeks				Y										
Biomarkers (Part A)														
Change from baseline in biomarkers including serum C4, total serum bile acids, autotaxin, FGF-19, and IL-31				Y[1]				Y			Y[1]	Y		
Biomarkers (Part B)														
Change from baseline in biomarkers including serum C4, total serum bile acids, autotaxin, FGF-19, and IL-31				Y[1]							Y[1]	Y		
Effect of BABR on linerixibat PD (Part A)														
Biomarkers including serum C4, total serum bile acids, autotaxin, FGF-19, IL-31				Y[2]				Y[2]			Y[2]			
Effect of BABR on linerixibat PD (Part B)														
Biomarkers including serum C4, total serum bile acids, autotaxin, FGF-19, IL-31				Y[2]							Y[2]			
PK (Part A)														
PK concentration[3]				Y[4,5]	Y[4]									
Marker of liver fibrosis (Part A)														
Change from baseline in ELF at Week 24				Y							Y			
Change from baseline in Fibroscan at Week 24				Y							Y			
Marker of liver fibrosis (Part B)														
Change from baseline in ELF at Week 8				Y							Y			
Symptoms and quality of life (Part A)														
Change from baseline in ESS over 24 weeks				Y				Y			Y			
Change from baseline in BDI-II at 24 weeks				Y				Y			Y			

Endpoints	Absolute							Change from Baseline						
	Stats Analysis			Summary		Individual		Stats Analysis			Summary		Individual	
	T	F	L	T	F	F	L	T	F	L	T	F	F	L
Change from baseline in MFS over 24 weeks				Y			Y	Y			Y	Y		
Symptoms and quality of life (Part B)														
Change from baseline in ESS over 8 weeks				Y							Y			
Change from baseline in BDI-II at 8 weeks				Y							Y			
Change from baseline in MFS over 8 weeks				Y			Y				Y			
Symptoms and quality of life (Part A and Part B)														
Change from baseline in WIS							Y					Y		
Change from baseline in WSS							Y					Y		
Change from baseline in WFS							Y					Y		

NOTES :

- T = Table, F = Figures, L = Listings, Y = Yes display generated
- Stats Analysis = Represents TFL related to any formal statistical analyses (i.e. modelling) conducted.
- Summary = Represents TFL related to any summaries (i.e. descriptive statistics) of the observed raw data.
- Individual = Represents FL related to any displays of individual participant observed raw or derived data.

1. Descriptive summary tables will also be provided by the following covariates: Baseline cholestyramine use; Baseline concomitant UDCA, and Baseline concomitant cholestatic pruritus therapy.
2. The descriptive analysis of these endpoints will be by baseline BABR. See the baseline BABR definition in Section 4.6.1.
3. Population PK analysis will be described in a separate report.
4. PK population will be used in the analysis.
5. Besides summarizing PK concentration by visit, PK concentration will also be summarized by baseline BABR (see Section 4.6.1 for details) and visit as well as presence or absence of cirrhosis (as collected on PBC history CRF) and visit, and baseline CKD category and visit (see Section 6.3.6.2 for details) in separate tables.

4.4.2. Statistical Analyses of Exploratory Endpoints

- Change from baseline in concentration of exploratory biomarkers

The change from baseline in exploratory biomarkers at visits 4 – 9 will be fitted to MMRM. The comparison of effects averaged over the visits will be estimated from the model.

Endpoint / Variables
• Change from baseline in concentrations of exploratory biomarkers including: ○ Serum C4 ○ Total serum bile acids ○ Autotaxin ○ Fibroblast Growth Factor-19 (FGF-19) ○ Interleukin 31 (IL-31)
Data Transformation
• The values of biomarkers will be natural log transformed to obtain log-transformed values at baseline, visits 4 – 9 of the study intervention during Part A. • The estimated change from baseline in log-transformed value will be transformed back to obtain geometric ratio to baseline.

- The estimated difference vs placebo in log-transformed value will be transformed back to obtain geometric ratio to placebo.
- The % coefficient of variation of estimate will be calculated as

$$\%CV_w = 100 * \sqrt{\exp(SE^2) - 1}$$

Model Specification

- Mixed Models Repeated Measures (MMRM) analysis.
- Terms in the model:

$$Y = Treatment + Visit + Treatment * Visit + BL_{Biomarker} + BL_{Biomarker} * Visit + BL_{Concomitant\ itch\ med} + Error$$

  - **Response** (Y): change from baseline in log transformed biomarker data at visits 4 - 9 of the study intervention during Part A.
  - **Categorical**: treatment group, visit, baseline concomitant itch therapy (3-level)
  - **Continuous**: log transformed biomarker baseline value
  - **Interaction**: treatment group*visit, log transformed baseline biomarker*visit
  - **Repeated**: visit
- Baseline is defined in Section 4.1.2

Model Checking & Diagnostics

- See Section 4.1.1

Model Results Presentation

- The back transformed LS means and the corresponding 95% confidence intervals of geometric ratio to baseline in biomarkers over 24 weeks for linerixibat and placebo will be reported. The back transformed LS mean and 95% CI for the geometric ratio to placebo will be reported in the same table.
- The back transformed LS means and 95% confidence intervals for geometric ratio to baseline in biomarkers at each visit (Weeks 4, 8, 12, 16, 20, and 24) will be reported for linerixibat and placebo, as well as the back transformed LS means and corresponding 95% CIs for the geometric ratio to placebo. A figure will be presented for the geometric ratio to placebo at each visit.

Subgroup/covariate Analyses

- The MMRM analysis will be repeated by including the covariates listed below and its interaction with treatment group:
  - Baseline cholestyramine use
  - Baseline concomitant UDCA
  - Baseline BABR

  Please refer to Section 4.6.1 on the details of covariates. When covariate baseline cholestyramine or baseline BABR use is included, baseline concomitant itch therapy will not be included in the model.
- The back transformed geometric ratio to baseline and its corresponding 95% CI will be reported for each treatment group, as well as the back transformed geometric ratio to placebo and its 95% CI. The p-values for treatment group, covariate, and interaction between treatment group and covariate will be reported.

- Change from baseline in ESS over 24 weeks

Change from baseline in ESS at visits 6 and 9 will be fitted by MMRM. The comparison of effects averaged over visits 6 and 9 will be estimated from the MMRM model.

Endpoint / Variables

- Change from Baseline in ESS over 24 weeks

Model Specification

- Mixed Models Repeated Measures (MMRM) analysis.
- Terms in the model:

$$Y = Treatment + Visit + Treatment * Visit + BL_{ESS} + BL_{ESS} * Visit + BL_{Concomitant\ itch\ med} + Error$$

  - **Response** (Y): change from baseline in ESS at visits 6 and 9 of the study intervention during Part A.

○ **Categorical**: treatment group, visit, baseline concomitant itch therapy (3-level) ○ **Continuous**: baseline ESS ○ **Interaction**: treatment group*visit, baseline ESS *visit ○ **Repeated**: visit • Baseline is defined in Section 4.1.2
Model Checking & Diagnostics
• See Section 4.1.1
Model Results Presentation
• The LS means and the corresponding 95% confidence intervals of change from baseline in ESS over 24 weeks for linerixibat and placebo, as well as the LS mean and 95% CI for the difference from placebo will be reported.

- Change from baseline in BDI-II at 24 weeks

Change from baseline in BDI-II at Visits 6 and 9 will be fitted by MMRM. The comparison at Visit 9 (week 24) will be estimated from the MMRM model.

Endpoint / Variables
• Change from Baseline in BDI-II at 24 weeks
Model Specification
• Mixed Models Repeated Measures (MMRM) analysis. • Terms in the model: $$Y = Treatment + Visit + Treatment * Visit + BL_{BDI-II} + BL_{BDI-II} * Visit + BL_{Concomitant\ itch\ med} + Error$$ ○ **Response** (Y): change from baseline in BDI-II at visits 6 and 9 of the study intervention during Part A. ○ **Categorical**: treatment group, visit, baseline concomitant itch therapy (3-level) ○ **Continuous**: baseline BDI-II ○ **Interaction**: treatment group*visit, baseline BDI-II *visit ○ **Repeated**: visit • Baseline is defined in Section 4.1.2
Model Checking & Diagnostics
• See Section 4.1.1
Model Results Presentation
• The LS means and the corresponding 95% confidence intervals of change from baseline in BDI-II at Week 24 for linerixibat and placebo, as well as the LS mean and 95% CI for the difference from placebo will be reported.

- Change from baseline in MFS as measured by 0 – 10 NRS over 24 weeks

Change from baseline in MFS at visits 4 - 9 will be fitted by MMRM. The comparison of effects averaged over the visits will be estimated from the MMRM model.

Endpoint / Variables
• Change from Baseline in MFS over 24 weeks

Model Specification
• Mixed Models Repeated Measures (MMRM) analysis. • Terms in the model: $$Y = Treatment + Visit + Treatment * Visit + BL_{MFS} + BL_{MFS} * Visit + BL_{Concomitant\ itch\ med} + Error$$ ○ **Response** (Y): change from baseline in MFS at visits 4 - 9 of the study intervention during Part A. ○ **Categorical**: treatment group, visit, baseline concomitant itch therapy (3-level) ○ **Continuous**: baseline MFS ○ **Interaction**: treatment group*visit, baseline MFS *visit ○ **Repeated**: visit • Baseline is defined in Section 4.1.2
Model Checking & Diagnostics
• See Section 4.1.1
Model Results Presentation
• The LS means and the corresponding 95% confidence intervals of change from baseline in MFS over 24 weeks for linerixibat and placebo, as well as the LS mean and 95% CI for the difference from placebo will be reported.

4.5. Safety Analyses

The safety analyses will be based on the Safety Analysis Set, unless otherwise specified. The summary of data during study intervention will be presented by Part A and Part B separately, unless otherwise specified. For all the adverse events analyses where cumulative incidence proportions are summarized, the exposure adjusted incidence rates will also be provided.

4.5.1. Extent of Exposure

The number of days exposure to study drug during the study intervention period and the cumulative dose will be summarized.

The formulas of calculation and the missing data rules are detailed in Section 6.3.8.1.

4.5.2. Adverse Events

Adverse event analyses, including adverse events (AEs), serious adverse events (SAEs) and adverse events of special interest (AESIs), will be based on GSK core data standards. Adverse events will be mapped to system organ classes (SOC) and preferred terms (PT) using the standard Medical Dictionary for Regulatory Affairs (MedDRA dictionary). The most recent MedDRA dictionary at the time of the primary analysis will be used. All AE and SAE summaries will be by SOC and PT unless otherwise specified. Liver related AEs will also be summarized by standardized MedDRA Queries (SMQs) and PT.

An adverse event (AE) is considered study intervention emergent if the AE onset date is on or after study intervention start date unless otherwise specified (see Section 6.3.2). All AE summaries will be based on study intervention emergent events unless otherwise specified. AEs and SAEs will be summarized by Parts A and B separately. If an AE/SAE

starts in Part A and lasts to Part B, it will be counted only in Part A. SAE summaries will be based on all SAEs reported regardless of whether they meet the definition of study intervention emergent or not.

For all AE tables, counting will be by participant and not by event. In other words, if a participant has more than one event with the same preferred term, the participant will be counted once for that preferred term in the treatment group. If a participant has more than one event in the same system organ class, the participant will be counted once for that system organ class in the treatment group. All AE and SAE summaries will be exposure adjusted unless otherwise specified.

AEs that are study intervention emergent will be summarized including:

- counts and percentages of participants with any AE,
- common AEs defined as AEs with ≥ 5% incidence in any treatment group by PT,
- AEs by maximum severity or grade (not exposure adjusted),
- AEs related to study treatment,
- AEs leading to permanent discontinuation of study intervention,
- AEs leading to study withdrawal

All AEs at Part A will also be summarized by operating model.

Common AEs defined as AEs with ≥ 5% incidence in any treatment group, will also be reported by PT for subgroups sex, age, race, and ethnicity, as defined in Section 4.6.1.

Under the supplementary safety estimand strategy, all AEs, will also be reported including AE's that occur after permanent treatment discontinuation.

SAEs during on-treatment periods will be summarized including:

- counts and percentages of participants with any SAE,
- SAE related to study treatment,
- fatal SAEs,
- fatal SAEs related to study treatment,
- number of participants and occurrences of SAE (not exposure adjusted)

Under the supplementary safety estimand strategy, SAEs will also be reported including SAE's that occur after permanent treatment discontinuation.

A study treatment-related AE is defined as an AE for which the investigator classifies the possible relationship to study treatment as "Yes". A worst-case scenario approach will be taken to handle missing relatedness data, i.e. the summary table will include events with the relationship to study treatment as 'Yes' or missing.

Complete participant listings of all AEs will be provided. For each AE the following will be specified: treatment group, start and stop dates, time since study treatment 1st and last

dose, severity grade, MedDRA system organ class and preferred term, duration, relationship to study treatment, action taken, outcome of the adverse event and seriousness. The Common Terminology Criteria for Adverse Events (CTCAE), currently in version 5, will be used to quantitate the severity of diarrhea as well as the standard AE intensity grading defined in protocol (TMF-16740250) Section 10.3.4. The severity evaluated by both criteria will be provided.

4.5.2.1. Adverse Events of Special Interest

The following will be considered adverse events of special interest (AESI) for the purpose of analyses:

- Diarrhea reported as an AE
  - Definition: all study intervention emergent AEs with High Level Term (HLT) of "Diarrhoea (excl infective)" that have been marked as AESI by investigators
- Elevated ALT reported as a Liver Event (meets Liver Monitoring or Stopping criteria)
  - Definition: study intervention emergent AEs with System Organ Class (SOC) of "Hepatobiliary disorders" or "Investigations" that have been marked as AESI by investigators

The summary of event characteristics will be provided for each AESI respectively, including:

- count and percentage of participants with any event,
- number of events,
- count and percentage of participants with any event that is serious,
- count and percentage of participants with any event that is related to study treatment,
- count and percentage of participants by number of occurrences (one, two, three or more), maximum severity or grade, maximum severity or grade for events related to study intervention, outcomes and the action taken for the event,
- count and percentage of events requiring antidiarrheal therapies and the number of cases resolved by the antidiarrheal therapies.

Under the supplementary safety estimand strategy, AESI will also be reported including AESI that occur after permanent treatment discontinuation.

The worst-case approach will be applied at a participant level for the maximum severity, i.e. a participant will only be counted once as the worst case from all the events experienced by the participant. For action taken to an event, a participant will be counted once under each action, e.g. if a participant has an event leading to both study intervention discontinuation and study withdrawal, the participants will be counted once under both actions. The maximum severity of diarrhea will be evaluated by both CTCAE and standard AE intensity grade and reported separately.

A Kaplan-Meier plot of time to first event of diarrhea reported as an AESI will be provided for Part B (a Kaplan-Meier plot of Part A data will be reported as part of the Integrated Safety Summary (ISS) and is described in the Project Data Analysis Plan for the ISS).

4.5.2.2. COVID-19 Assessment and COVID-19 AEs

The overall incidence of AEs and SAEs of COVID-19, COVID-19 AEs leading to study intervention discontinuation, COVID-19 AEs leading to study intervention interruption or delay will be summarized. The incidence of these events at individual PT level can be obtained from the standard AE/SAE summaries.

4.5.3. Additional Safety Assessments (if applicable)

4.5.3.1. Laboratory Data

Summaries of worst-case changes from baseline with respect to potential clinical importance (PCI) range will be generated for haematology, coagulation, fat soluble vitamins, and chemistry (including liver function tests) laboratory tests, where a PCI range is specified in Section 6.3.1. Decreases to low, changes to within range or no changes from baseline, and increases to high will be summarized for the worst-case post-baseline. If a participant has a decrease to low and an increase to high during the same time interval, then the participant is counted in both the “to Low” categories and the “to High” categories.

The observed and change from baseline of haematology, coagulation, fat soluble vitamins, and chemistry laboratory tests will be produced by visit and by treatment group. Liver function laboratory tests will be included with chemistry lab tests.

For haematology, coagulation, fat soluble vitamins, and chemistry (including liver function tests) laboratory tests, observed values will be compared to normal ranges and classified as “Low” (where applicable), “Normal”, or “High (if applicable). For each liver function test, 4x4 contingency tables (shift tables) will be displayed by treatment group and by visit that summarize the counts of participants who were classified as “Low”, “Normal”, “High” at Baseline vs. “Low”, “Normal”, “High” at Post-baseline time points, as well as the total. Worst-case post-baseline will also be reported. Summary and listing of abnormal laboratory results will be provided.

An e-DISH plot of maximum post baseline total bilirubin versus maximum post baseline ALT will be created for each intervention period separately (Part A and Part B). Similarly, e-DISH plots of maximum post baseline total bilirubin versus maximum post baseline AST, and maximum post baseline total bilirubin versus maximum post baseline ALP, will also be created for each study period.

For each of the respective plots, the following datapoints will be flagged:

- Maximum postbaseline total bilirubin equal to or exceeding 2 x ULN within 30 days after a maximum postbaseline ALT equal to or exceeding 3 x ULN
- Maximum postbaseline total bilirubin equal to or exceeding 2 x ULN within 30 days after a maximum postbaseline AST equal to or exceeding 3 x ULN
- Maximum postbaseline total bilirubin equal to or exceeding 2 x ULN within 30 days after postbaseline ALP became equal to or exceeding 2 x ULN

Plots will be generated using central data only and repeated using central plus local lab data.

The count and percentage of participants who have met the liver increased monitoring criteria and participants who have met the liver stopping criteria as defined in the protocol will be summarized for each treatment group.

4.5.3.2. Vital Signs

Change from baseline in vital signs, including temperature, pulse rate (bpm), systolic blood pressure(mmHg) and diastolic blood pressure (mmHg), will be summarized by visit and by treatment group. Summary of worst-case change from baseline with respect to potential clinical importance range will be produced.

4.5.3.3. ECG

The QTc data analysis will use the collected values based on Bazett's formula (QTcB), Fridericia's formula (QTcF), and/or another method. Change from baseline in ECG will be summarized by visit and by treatment group.

The QTc values based on Fridericia formula or Bazett's formula will be rounded to the nearest integer and the values will be categorized into the following CTCAE grade and ranges: Grade 0 (<450 msec), Grade 1 (450-480 msec), Grade 2 (481-500 msec), and Grade 3 (≥501 msec). Summaries of grade increase will be provided. These summaries will display the number and percentage of participants with any grade increase, increase to grade 2 and increase to grade 3 for the worst-case post-baseline observation only. Missing baseline grade will be assumed as grade 0.

The changes in QTc values will be categorized into the clinical concern ranges which are specific to changes in QTc: ≤30, 31-60, and >60 msec. A summary of change in QTc value will display the number and percentage of participants with a change within each range for the worst-case post-baseline observation only. Participants with missing baseline values will be excluded from this summary.

4.5.3.4. Fecal Occult Blood Test (FOBT)

Summary of count and percentage of participants with positive and negative FOBT at V3 and V9 and a shift table will be reported.

4.6. Other Analyses

4.6.1. Covariate/Subgroup analyses

The list of covariates and other strata may be used in the statistical analyses of primary endpoint are shown below. Additional covariates and other strata of clinical interest may also be considered.

Table 3 List of covariate/subgroup

Covariate/Subgroup	Details
Baseline cholestyramine use	Previous cholestyramine use (no concomitant use at baseline) Baseline concomitant cholestyramine Baseline cholestyramine naive
Region	North America (US and Canada) Asia (China and Japan) Europe (Belgium, Bulgaria, Czechia, France, Germany, Greece, Italy, Israel, Poland, Spain, Switzerland, United Kingdom) Other Regions (Argentina, Brazil, Mexico, Russian Federation)
Baseline ALP*	Continuous
Baseline ALP* (Categorical)	≤ 1xULN > 1xULN to < 1.67xULN ≥ 1.67xULN
Baseline total bilirubin*	Continuous
Baseline total bilirubin (Categorical)*	≤ 1xULN > 1xULN
Baseline concomitant UDCA	Yes No
Baseline total serum bile acids	Continuous
Baseline itch severity group*	≥4 - <7 'Moderate' ≥7 - ≤10 'Severe'
Baseline MIS*	Continuous
Baseline concomitant cholestatic pruritus therapy (Yes vs. No)	Yes No Note: if a participant used any of the listed concomitant background itch therapy then yes, otherwise it is no.
Baseline concomitant cholestatic pruritus therapy (3-level)	Regimen contains BABR, including cholestyramine, colesevalem, colestimide and colestipol Regimen that does not contain BABR, including antihistamines#, bezafibrate, fenofibrate, rifampicin, naltrexone, naloxone, nalfurafine, gabapentin or sertraline No defined anti-cholestatic pruritus treatment. #Note: When antihistamines are used acutely for indications other than itch (e.g. for treatment of acute allergic reactions), it is not considered as an anti-cholestatic pruritus treatment, therefore, it is not considered as a regimen that doesn't contain BABR. Please refer to protocol Section 6.9.1 for the list of anti-cholestatic pruritus treatment.
Baseline concomitant Fibrates (Yes vs. No)	Yes No

Covariate/Subgroup	Details
	Note: if a participant used any Fibrates then yes, otherwise it is no.
Baseline concomitant BABR (Yes vs. No)	Yes No Note: if a participant used any of the BABR then yes, otherwise it is no.
Sex	Female Male
Age Category (years)	18-49 50-64 ≥65
Race	Asian White Other
Ethnicity	Not Hispanic or Latino and Other Hispanic or Latino

*Covariates will be included in the summary of baseline covariates only and will not be included in the subgroup analysis detailed in Section 4.2.2.

4.6.2. Subpopulation Analyses

4.6.2.1. Analysis Sets for Subpopulation Analyses

The following analysis sets are defined for subpopulation analyses to support regional regulatory submission. Additional analysis sets may be considered if requested by regulators.

Analysis Set	Definition / Criteria	Analyses Evaluated
Screening China	• This is a subset of the Screening analysis set enrolled at sites in China Mainland	• Study Population
Enrolled China	• This is a subset of the Enrolled analysis set for participants enrolled at sites in China Mainland	• Study Population
ITT China	• This is a subset of the ITT Population enrolled at sites in China Mainland	• Efficacy
Safety China	• This is a subset of the Safety analysis set enrolled at sites in China Mainland	• Safety
Non-BABR ITT China	• This is a subset of the Non-BABR ITT Population enrolled at sites in China Mainland	• Efficacy
PK China	• This is a subset of the PK analysis set enrolled at sites in China Mainland	• PK

Analysis Set	Definition / Criteria	Analyses Evaluated
Screening Japan	• This is a subset of the Screening analysis set enrolled at sites in Japan whose race is Japanese	• Study Population
Enrolled Japan	• This is a subset of the Enrolled analysis set for participants enrolled at sites in Japan whose race is Japanese	• Study Population
ITT Japan	• This is a subset of the ITT Population enrolled at sites in Japan whose race is Japanese.	• Efficacy
Safety Japan	• This is a subset of the Safety analysis set enrolled at sites in Japan whose race is Japanese	• Safety
PK Japan	• This is a subset of the PK analysis set enrolled at sites in Japan whose race is Japanese	• PK
Screening East Asia	• This is a subset of the Screening analysis set who are of East Asia subpopulation*	• Study Population
Enrolled East Asia	• This is a subset of the Enrolled analysis set who are of East Asia subpopulation*	• Study Population
ITT East Asia	• This is a subset of the ITT Population who are of East Asia subpopulation*	• Efficacy
Safety East Asia	• This is a subset of the Safety analysis set who are of East Asia subpopulation*	• Safety
Non-BABR ITT East Asia	• This is a subset of the Non-BABR ITT Population who are of East Asia subpopulation*	• Efficacy
PK East Asia	• This is a subset of the PK analysis set who are of East Asia subpopulation*	• PK

* East Asia subpopulation: participants of a relevant Asian heritage (Asian – Japanese Heritage, Asian – East Asian Heritage) enrolled at sites in Japan, China mainland, Hong Kong, Taiwan, Macau, South Korea.

4.6.2.2. China Subpopulation Analyses

The analyses of China subgroup are based on ITT China, Safety China and PK China analysis sets. The following endpoints will be evaluated:

	Endpoints
Primary	• Change from Baseline in Monthly Itch Scores over 24 weeks using a 0-10 numerical rating scale (NRS)
Secondary	• Change from baseline in Weekly Itch score at Week 2 • Change from Baseline in Monthly Sleep Score as measured by 0-10 NRS over 24 weeks • Responder defined as achieving a ≥2-point reduction from Baseline in the Monthly Itch score at Week 24. • Responder defined as achieving a ≥3-point reduction from Baseline in the Monthly Itch score at Week 24. • Responder defined as achieving a ≥4-point reduction from Baseline in the Monthly Itch score at Week 24. • Change from Baseline in PBC-40 domain scores at Week 24
Safety	• Adverse Events (AEs) and Serious Adverse Events (SAEs) • Clinical laboratory evaluation
Exploratory	• Change from Baseline in concentrations of exploratory biomarkers including: o Serum C4 o Total serum bile acids o Autotaxin o Fibroblast Growth Factor-19 (FGF-19) • PK

The analyses of these endpoints will be conditional on the primary endpoint results for all participants. If the primary endpoint for all participants is not significant ($p > 0.05$), the China subgroup analyses may not be produced.

For responder analysis, if the sample size per stratum permits, CMH estimation will be applied. If the number of participants per stratum is small, an unadjusted estimation will be used. For primary and the other secondary endpoints, if data permits, the same MMRM as the main analytic approach will be applied to the China subpopulation using the same covariates. When there is statistical model involved, LS means and 95% CI will be reported, whereas p-value will not be provided.

4.6.2.3. Japan Subpopulation Analyses

The analyses of Japan subgroup are based on ITT Japan, Safety Japan and PK Japan analysis sets. The following endpoints will be evaluated:

	Endpoints
Primary	• Change from Baseline in Monthly Itch Scores over 24 weeks using a 0-10 numerical rating scale (NRS)
Secondary	• Change from baseline in Weekly Itch score at Week 2 • Change from Baseline in Monthly Sleep Score as measured by 0-10 NRS over 24 weeks • Responder defined as achieving a ≥2-point reduction from Baseline in the Monthly Itch score at Week 24. • Responder defined as achieving a ≥3-point reduction from Baseline in the Monthly Itch score at Week 24. • Responder defined as achieving a ≥4-point reduction from Baseline in the Monthly Itch score at Week 24. • Change from baseline in Patient's Global Impression of Severity (PGI-S) over 24 weeks • PGI-C over 24 weeks • Change from baseline in ALP at Week 24 • Change from baseline in bilirubin at Week 24
Safety	• Adverse Events (AEs) and Serious Adverse Events (SAEs) • Clinical laboratory evaluation (including liver chemistry panel and fasting lipids)
Exploratory	• Change from Baseline in concentrations of exploratory biomarkers including: o Serum C4 o Total serum bile acids o Autotaxin o Fibroblast Growth Factor-19 (FGF-19) o Interleukin 31 (IL-31) • PK

The analyses of these endpoints will be conditional on the primary endpoint for all participants. If the primary endpoint for all participants is not significant ($p>0.05$), the Japan subgroup analyses may not be produced.

For responder analysis, if the sample size per stratum permits, CMH estimation will be applied. If the number of participants per stratum is small, an unadjusted estimation will be used. For primary and the other secondary endpoints, if data permits, the same MMRM as the main analytic approach will be applied to the Japan subpopulation using the same covariates. When there is a statistical model involved, LS means and 95% CI will be reported, whereas p-value will not be provided.

4.6.2.4. East Asia Subpopulation Analyses

The analyses of East Asia subgroup are based on ITT East Asia, Safety East Asia and PK East Asia analysis sets. The following endpoints will be evaluated:

	Endpoints
Primary	• Change from Baseline in Monthly Itch Scores over 24 weeks using a 0-10 numerical rating scale (NRS)
Secondary	• Change from baseline in Weekly Itch score at Week 2 • Change from Baseline in Monthly Sleep Score as measured by 0-10 NRS over 24 weeks • Responder defined as achieving a ≥2-point reduction from Baseline in the Monthly Itch score at Week 24. • Responder defined as achieving a ≥3-point reduction from Baseline in the Monthly Itch score at Week 24. • Responder defined as achieving a ≥4-point reduction from Baseline in the Monthly Itch score at Week 24. • Change from Baseline in PBC-40 domain scores at Week 24
Safety	• Adverse Events (AEs) and Serious Adverse Events (SAEs) • Clinical laboratory evaluation
Exploratory	• PK

The analyses of these endpoints will be conditional on the results of primary endpoint for all participants. If the primary endpoint for all participant is not significant ($p > 0.05$), the East Asia subgroup analyses may not be produced.

For responder analysis, if the sample size per stratum permits, CMH estimation will be applied. If the number of participants per stratum is small, an unadjusted estimation will be used. For primary and the other secondary endpoints, if data permits, the same MMRM as the main analytic approach will be applied to the East Asia subpopulation using the same covariates. When there is statistical model involved, LS means and 95% CI will be reported, whereas p-value will not be provided.

4.6.3. Operating model analyses

There will be 3 operating models in this study:

- Metasite model by Metasite (portion of US participants, currently none planned outside US)
- Flexible model by brick & mortar site (global)
- Brick & mortar model by brick & mortar site (global)

The following definitions will be used to categorise participants by operating models:

- Metasite model is defined as having all visits remotely for participants at Metasite.
- Flexible model is defined as having at least one remote visit during the study intervention period of Part A for participants at brick & mortar sites.
- Brick and mortar model is defined as having all visits onsite during the study intervention period of Part A for participants at brick & mortar sites.

Visits conducted remotely will be captured via the eCRF at each visit using the question "visit conducted remotely" (Yes/No). This will be used to define the operating model for each participant as described above.

The following summaries will be reported for Part A:

- Participant disposition by site operating model (in addition to overall)
- Important protocol deviations by site operating model (in addition to overall)
- Adverse events by site operating model (in addition to overall)

To assess the impact of the three operating models (Metasite vs Flexible vs Brick and Mortar) on the primary endpoint, the following MMRM analysis will be applied:

- **Response**: change from baseline in MIS at Week 4, 8, 12, 16, 20, and 24 of the study intervention during Part A.
- **Categorical variables**: treatment group, visit (Weeks 4, 8, 12, 16, 20, and 24), baseline concomitant therapy, country/region (US or RoW), operating model (Metasite, Flexible, or Brick and Mortar)
- **Continuous variable**: baseline MIS
- **Interaction terms**: treatment group*visit, treatment group*country/region, treatment group*operating model, baseline MIS*visit,
- **Repeated**: visit

In addition to the p-values of operating model and the interaction between treatment and operating model, the following estimates will be reported:

- LS means and the corresponding 95% CI of change from baseline on MIS over 24 weeks for each treatment group and between treatment group comparisons for Metasite vs Flexible vs Brick and Mortar operating models.
- The same MMRM analysis will be applied to the participants in US region only (country/region terms will therefore be removed from the model). LS means of change from baseline over 24 weeks for each treatment group and between treatment group comparisons for Metasite vs Flexible vs Brick and Mortar will be reported, based on US region only.

4.6.4. Shrinkage Analysis

Bayesian hierarchical modelling will be used to estimate treatment effect estimates for the subgroups in Table 3 of Section 4.6.1, as mentioned in Section 4.2.2.

The primary endpoint in each subgroup-level will first be analyzed using the same method as per Section 4.2.2. For each subgrouping variable, each subgroup-level specific least square mean treatment difference estimates and their corresponding standard error will then be included in a Bayesian hierarchical model to obtain shrinkage estimates of the subgroup-specific treatment effects.

The subgroup-specific treatment effects, μ_j ($j = 1, \ldots, J$ where J is the number of subgroups for a particular subgrouping variable) will be assumed to be normally distributed as $\mu_j \sim N(\mu, \tau^2)$ where μ is the overall treatment effect and τ^2 is the between-subgroup variance. The following non-informative prior for μ and weakly-informative prior for τ will be used in the model:

a. $\mu \sim$ U[-16, 16]
b. $\tau \sim$half-$N(SD{=}1)$

The selected prior for τ covers a very small to very large between subgroup heterogeneity given the assumed SD for the primary endpoint of 2.07 [Wang, 2024].

Models will be fitted using a Markov chain Monte Carlo (MCMC) algorithm with at least 50,000 iterations, a thin of at least 5 and a burn-in of at least 1000. The posterior mean differences (μ_j) with 95% credible interval (CrI) will be presented by subgroup in a forest plot alongside the initial estimates. A posterior summary of τ will also be presented (median and CrI).

4.7. Interim Analyses

An interim analysis will be performed after approximately 100 participants reach 24 weeks of treatment or discontinue early. The interim analysis will be evaluated by an Independent Monitoring Committee (IDMC) and conducted such that the ongoing study integrity is maintained. Only the independent statistical support group responsible for providing the interim analysis results to the IDMC will be unblinded to the individual treatment group assignments (current version: GlaxoSmithKline Document Number TMF-16125967). Individual and detailed interim analysis results will not be shared with investigators, participants, or the study team who are involved in the conduct of the study before the final database lock.

4.7.1. Interim Analysis Set

The analysis sets used in the interim analysis are defined as below:

Analysis Set	Definition / Criteria	Analyses Evaluated
ITT Interim	• All randomized participants through the data cut off of interim analysis. Participants will be classified according to the treatment as randomized.	• Study Population • Efficacy
ITT Interim Primary	• The first approximately 100 randomized ITT participants who have completed 24 weeks of treatment or who have discontinued treatment early.	• Efficacy
Safety Interim	• All randomized participants through the data cut off of interim analysis who take at least 1 dose of study intervention. Participants will be analyzed according to the treatment as randomized.	• Study Population • Safety

4.7.2. Interim Study Population Analyses

Interim study population analyses will be based on ITT Interim and Safety Interim analysis sets.

Summary of disposition, including study withdrawal and treatment discontinuation, will be provided. Demographic characteristics, including sex, age, age group, ethnicity, race, will be summarized. Summaries of the covariate/subgroups at baseline will be provided. See Section 4.6.1 for the full list of covariate/subgroups.

Summary of disposition, including study withdrawal and treatment discontinuation will also be provided based on the ITT Interim Primary analysis set.

4.7.3. Interim Efficacy Analyses

4.7.3.1. Primary analyses

The Interim analysis for the primary efficacy endpoint will use the ITT Interim Primary analysis set, using MMRM as described in Section 4.2.

The LS means and 95% confidence intervals will be reported for change from baseline in MIS over 24 weeks. The LS mean, 95% CI, and p-value will be reported for the difference between placebo and linerixibat in change from baseline in MIS over 24 weeks. The predictive probability of success with complete data given interim data, and the conditional power given interim data will be provided.

The futility assessment, the calculation of predictive probability of success and conditional power are detailed in the IDMC charter (TMF-16125967)Appendix 3: Statistical Guidance and Stopping Rule – Efficacy.

4.7.3.2. Supplementary analyses

The supplementary analyses of efficacy will be based on ITT Interim analysis set.

- MIS

The LS means and 95% CIs of change from baseline in MIS and difference from placebo over 24 weeks will be reported. The LS means and 95% CI of change from baseline in MIS and difference from placebo at each visit (Weeks 4, 8, 12, 16, 20, 24 since Day 1 of active treatment) from MMRM analysis described in Section 4.2 will also be reported.

- MSS

The MMRM analysis for the sleep endpoint described in Section 4.3.1 will be used for the interim analysis of MSS. The LS means and 95% CIs of change from baseline in MSS and difference from placebo over 24 weeks will be reported. The LS means and 95% CIs of change from baseline in MSS and difference from placebo at each visit (Weeks 4, 8, 12, 16, 20, 24 since Day 1 of active treatment) will also be reported.

4.7.4. Interim Safety Analyses

The safety analyses at interim will be based on Safety Interim analysis set.

Summaries of AEs of diarrhea or elevated ALT leading to permanent study drug discontinuation will be provided. The liver function test results (i.e., AST, ALT, total bilirubin, ALP, INR, and GGT) for participants meeting liver monitoring and/or stopping criteria will be listed, which will include participants with elevated ALT leading to permanent study drug discontinuation.

As supplementary analyses, summaries of all AEs and SAEs, summary of AEs leading to permanent study drug discontinuation, and summary of participants meeting liver monitoring and/or stopping criteria will also be provided.

4.8. Changes to Protocol Defined Analyses

Changes from the originally planned statistical analysis specified in the protocol are detailed in Table 4.

Table 4 Changes to Protocol Defined Analysis Plan

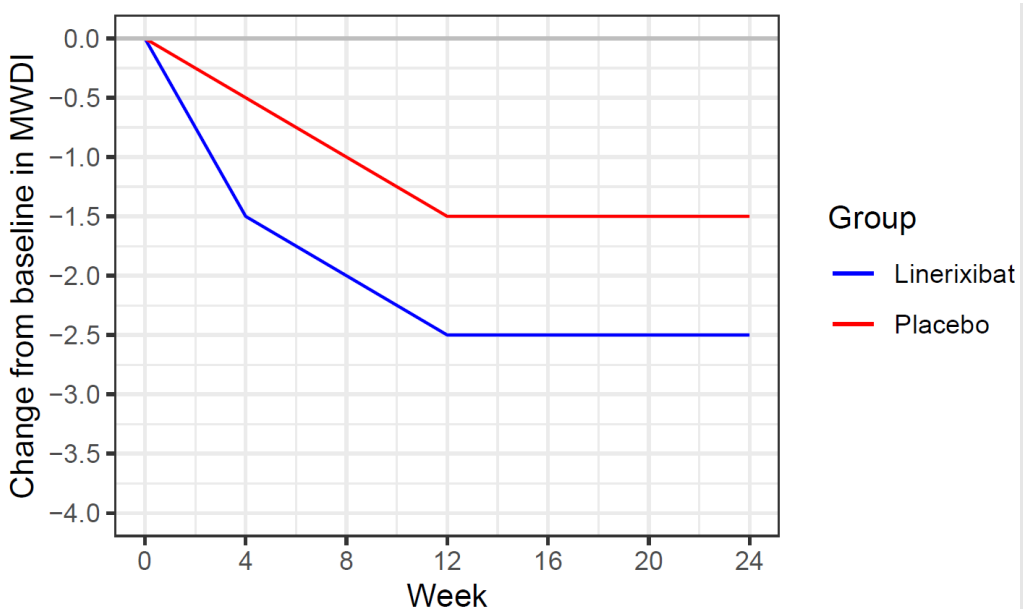
Protocol Defined Analysis	SAP Defined Analysis	Rationale for Changes
• Hypothetical strategy will be used for COVID-19 related intercurrent events for all primary and secondary endpoints.	• Hypothetical strategy will be used for intercurrent events caused by COVID-19 related site closure, lockdowns, or other supply chain issues for the primary and key secondary endpoints. Treatment policy will be used for the supportive secondary endpoints.	• In line with regulatory agency's evolving recommendation on handling of "COVID-19 related" intercurrent events since COVID-19 pandemic is no longer a temporary problem. Therefore, it is not necessary to apply this strategy to the supportive secondary endpoints.
• PGI-C will be analyzed as a continuous variable using a MMRM model.	• Additionally, the percentage of participants with improved PGI-C for each treatment group will be reported, as well as the percentage difference, 95% CI, and p-value for the difference from placebo.	• Analyzing PGI-C as a categorical variable may aid interpretation.
• ALP and Bilirubin will be analyzed using the ITT population only.	• Additionally, statistical analysis for ALP and Bilirubin will be repeating for the ALP ≥1.67x ULN ITT population (see Section 4.3.2).	• For demonstrating absence of worsening disease (PBC) progression.
• Determination of PK parameters by population PK analysis.	• Population PK analysis will no longer be performed. Only PK concentration data will be summarized in this study.	• Due to the minimal absorption of linerixibat (oral bioavailability = 0.05%), plasma concentrations are intermittently measurable with a high degree of variability. Therefore, a population PK analysis is not feasible.
• PRO compliance analysis not specified.	• PRO compliance Section 6.2 added. Summaries of compliance over all PROs will be performed and reported as a percentage of completed PROs out of the total number of expected PROs.	• In line with updates to internal guidance and statistical analysis plan template.

5. SAMPLE SIZE DETERMINATION

Approximately 230 participants (115 per group) who pass the screening eligibility criteria will be randomized in a 1:1 ratio to linerixibat or placebo groups. The sample size of 230 was obtained and evaluated assuming that 10% of participants discontinue from the study at random and the data from the discontinuation week through Week 24 are missing. This sample size allows approximately 90% power for a two-sided test with an alpha level of 0.05 when the expected overall difference in change from baseline in MIS across 24 weeks is ~-0.9 (based on BAT117213 and GLIMMER data) between the linerixibat and placebo groups (assuming a between subject SD of 2.07 for both groups base on GLIMMER data).

Change from baseline in the WIS was simulated for each week from Week 1 to Week 24 after randomization. It was assumed that the mean change from baseline in the WIS of placebo group reaches its plateaued effect size of -1.5 by week 12 in a linear relationship against treatment week. It was also assumed that delta between linerixibat and placebo reaches plateau of -1 by week 4 in a linear relationship against treatment week, and the mean change from baseline in the WIS of linerixibat group is the sum of placebo and delta as shown in Figure 1.

Figure 1 Expected Change from Baseline in MWDI*



*MWDI now referred to as the Weekly Itch Score (WIS).

The expected weekly means are listed below:

Table 5 Expected Change from Baseline in Weekly Itch Score (WIS)

Group	N	Expected Change from Baseline in WIS by Week (Week 1 – Week 11)	Expected Change from Baseline in WIS (Week 12 - 24)	Dropout Rate
Placebo	115	(-0.125, -0.250, -0.375, -0.500, -0.625, -0.750, -0.875, -1.000, -1.125, -1.250, -1.375)	-1.5	10%
Linerixibat	115	(-0.375, -0.750, -1.125, -1.500, -1.625, -1.750, -1.875, -2.000, -2.125, -2.250, -2.375)	-2.5	10%

For each iteration of simulation, change from baseline data for each week was generated for 115 subjects per group using a truncated multivariate normal distribution within the interval of [-10, 10] with the mean vectors presented above. The variance covariance matrix was assumed to have a compound symmetry (CS) structure with standard deviation of 2.25 and correlation of 0.85 which corresponds to a between subjects SD of ~2.07 and a within subject SD of ~0.87, based on the GLIMMER data.

The rates of study withdrawal that resulted in missing values from withdrawals for placebo group and linerixibat 40 mg BID group after randomization were 1/36 and 1/23, respectively, in the GLIMMER study. Therefore, it was expected to have ~10% study withdrawal rate for each arm and the study withdrawal can occur at any week after randomization. Missing data was assumed to be missing at random in the simulation.

Simulations were based on planned visits at Week 4, 8, 12, 16, 20, and 24. For visit on Week 4, the worst change from baseline in WIS from Week 1 to 4 was reported as the change from baseline in MIS. Similarly, the change from baseline in MIS were obtained for other visits. Mixed model repeated measure analysis was then applied to change from baseline in MIS and the model included treatment, visit, and treatment by visit interaction as fixed effect. The comparison of active and placebo arms was based on the simple average of LS means of all visits, i.e., the main effect of treatment averaged across the 6 visits with equal weight at each visit.

6. SUPPORTING DOCUMENTATION

6.1. Appendix 1 Study Population Analyses

Unless otherwise specified, study population analyses will be based on the ITT Population and Safety analysis sets and summarized by Part A and Part B.

A summary of the number of participants in each of the participant level analyses set will be provided. Enrolment will be presented by country and site based on the Enrolled analysis set. Screen status and reasons for screen failures will be based on Screening analysis set.

6.1.1. Participant Disposition

A summary of the number and percentage of participants who completed the study as well as those who prematurely withdrew from the study will be provided. Reasons for study withdrawal will be summarized. Participant status and participant disposition at Part A will also be summarized by site operating model (see Section 4.6.3).

6.1.2. Demographic and Baseline Characteristics

The demographic characteristics including age, sex, ethnicity, height/weight at screening and race will be summarized with descriptive statistics. In addition, the following age categories will be summarized: 18-49, 50-64, and ≥65; and < 65, 65-74 and ≥75. Additionally, a de-identified high-level race variable will be defined for the purpose of public disclosure as White, Asian, and Other (De-identified), where the categories of American Indian or Alaska Native/Black or African American/Mixed Race/Unknown/Not Reported/Missing are collapsed into the Other (de-identified) category. Similarly, if the summary of any demographics and/or that of race and racial combinations meet the criteria for de-identification, then a de-identified version should be produced.

A summary of baseline disease characteristics will be provided. This summary will include PBC, hepatitis, and cirrhosis diagnosis; liver biopsy; duration of pruritus; PBC historical stage; and prior liver transplant.

Past medical conditions and current medical conditions as of screening will be summarized respectively.

Substance use, including smoking history, tobacco use, alcohol and drug history will be summarized.

Number and percentage of participants taking UDCA, anti-cholestatic pruritus treatment, and antihistamines as anti-cholestatic pruritus treatment at baseline will be summarized.

6.1.3. Protocol Deviations

Important protocol deviations will be summarized.

Protocol deviations will be tracked by the study team throughout the conduct of the study. These protocol deviations will be reviewed to identify those considered as important as follows:

- Data will be reviewed prior to freezing the database to ensure all important deviations (where possible without knowing the study intervention details) are captured and categorised in the protocol deviations dataset.
- This dataset will be the basis for the summaries of important protocol deviations.

In addition to the overall summary of protocol deviations, separate summaries of

- important protocol deviations at Part A
- important protocol deviations at Part A by operating model will also be produced.

6.1.4. Prior and Concomitant Medications

Prior and concomitant medications will be coded using the WHO Drug dictionary (the most recent dictionary at the time of the primary analysis will be used). The summary of concomitant medications will be provided by ingredient, i.e. multi-ingredient medications will be summarized for each individual ingredient rather than a combination of ingredients. The summary will be created using ingredient base names, i.e. ingredients with the same base name but different salt will appear under one base name in the summary. Anatomical Therapeutic Chemical (ATC) classifications will appear in the summary.

The number and percentage of participants using rescue medication will be summarized by treatment group.

6.1.5. Study Intervention Compliance

Study intervention compliance (%) will be calculated as [Total cumulative actual dose / Total cumulative scheduled dose] *100 based on number of tablets dispensed and returned, as well as based on eDiary.

Overall compliance will be summarized using descriptive statistics as well as the categories of <80%, 80%-105%, and >105%. Listing of compliance will be provided.

6.2. Appendix 2 Electronic Patient-reported Outcome (ePRO) Compliance

Compliance is the state of being in accordance with protocol-defined specifications. Evaluating compliance with the schedule of PRO assessments is important to understand the completeness of the PRO data in this study. Study level compliance will be calculated to evaluate compliance across all PRO assessments in the study (i.e., assessments supporting primary, secondary, and exploratory endpoints), while endpoint level compliance will be a targeted calculation to evaluate compliance with PRO assessments supporting the analysis of primary and important secondary endpoints. Details for each of these calculations are provided below.

6.2.1. Study Level Compliance

Overall PRO compliance (across all PROs and all participants) for the study is calculated for the ITT population as:

$$\frac{\sum_{i=1}^{i=n} Number\ of\ complete\ PROs}{\sum_{i=1}^{i=n} Expected\ number\ of\ complete\ PROs},$$

Where i = participant and n = number of randomized participants.

Compliance will be calculated for each treatment group and each respective study period (Part A and Part B separately). The ‘Number of complete PROs’ will be calculated across

the study period for all participants in the ITT population. The definition for 'Expected number of complete PROs' is explained for each PRO below.

For the itch eDiary, 2 entries per participant per day are expected. For the sleep eDiary 1 entry per participant per day is expected, and for the fatigue eDiary 1 entry per participant per day is expected. For GSRS, 1 eDiary entry per participant per week is expected.

For visit-based PRO data, an assessment is expected for each participant at each visit identified in the SoA of the protocol (Section 1.3). For PRO assessments with multiple questions (PBC-40, ESS, BDI-II), a participant will be considered to have completed the questionnaire for each respective visit if at least one question from the questionnaire is completed. Participants who initiate completion of a PRO assessment at a site visit will generally complete the PRO in its entirety.

For Part A, a participant who completes Part A (including baseline) is anticipated to have the following PRO assessments:

- Itch eDiary; approximately 394 diaries* (2 diaries x from Day -28 to Day 169 or Visit 9 – 1, whichever is earlier)
- Sleep eDiary; approximately 197 diaries* (1 diary x from Day -28 to Day 169 or Visit 9 – 1, whichever is earlier)
- Fatigue eDiary; approximately 197 diaries* (1 diary x from Day -28 to Day 169 or Visit 9 – 1, whichever is earlier)
- GSRS; 25 eDiaries (baseline, plus 1 per week for 24 weeks)
- PGI-C; 6 entries (V4 to V9)
- PGI-S; 7 entries (V3 to V9)
- PBC-40; 7 entries (V3 to V9)
- BDI-II; 3 entries (V3, V6 and V9)
- ESS; 3 entries (V3, V6 and V9)

*for the purpose of this example the approximate number of eDiaries assumes that the subject had their Visit 9 on the planned day, but as per the formula the number of expected days is dependent on when Visit 9 actually occurs.

Therefore, a participant who does not withdraw from study before completion of Part A would have an anticipated number of complete PROs of 839 (813 diaries from eDiaries, and 26 from visit-based PRO assessments).

For Part B, a participant who completes Part B is anticipated to have the following PRO assessments:

- Itch eDiary; approximately 112 diaries* (2 diaries x from Visit 9 + 1 to Visit 9 + 56 or Visit 11 – 1, whichever is earlier)
- Sleep eDiary; approximately 56 diaries* (1 diary x from Visit 9 + 1 to Visit 9 + 56 or Visit 11 – 1, whichever is earlier)

- Fatigue eDiary; approximately 56 diaries* (1 diary x from Visit 9 + 1 to Visit 9 + 56 or Visit 11 – 1, whichever is earlier)
- GSRS; 8 eDiaries (1 per week for 8 weeks)
- PGI-C; 2 entries (V10 and V11)
- PGI-S; 2 entries (V10 and V11)
- PBC-40; 2 entries (V10 and V11)
- BDI-II; 1 entry (V11)
- ESS; 1 entry (V11)

*for the purpose of this example the approximate number of eDiaries assumes that the subject had their Visit 9 on the planned day, but as per the formula the number of expected days is dependent on when Visit 9 actually occurs.

Therefore, a participant who does not withdraw from study before completion of Part B would have an anticipated number of complete PROs of 240 (232 diaries from eDiaries, and 8 from visit-based PRO assessments).

If a participant is lost to follow-up, PRO compliance will be calculated up to the last Visit attended (i.e., anticipated number of eDiaries is dependent on the time they are in the study). Similarly, if a participant withdraws from the study, compliance should be calculated through to the early discontinuation/withdrawal (ED) visit.

The study level PRO compliance will be reported for each study period (Part A and Part B separately), for each treatment group and overall treatment groups. For Part A, the target compliance for eDiary assessments is 57% over all treatment groups (based on the requirement of only 4 daily scores per week, as described in Section 6.3.7). The target compliance for Part A visit-based PRO assessments is 80% over all treatment groups.

The study level PRO compliance calculation for Part B will be produced at the end of the study when all participants have either completed or withdrawn from the study.

6.2.2. Endpoint Level Compliance

For the primary endpoint in Part A (MIS), a summary table will be produced displaying the frequency and percentage of participants who have 0, 1, 2, 3, or 4 Weekly Itch Scores (WIS) available at each month. This will be presented by treatment group and overall treatment groups. Similarly, a summary table will also be produced for the key secondary endpoint related to MSS in Part A.

For the key secondary endpoint related to WIS in Part A, a summary table will be produced displaying the frequency and percentage of participants who have ≥4 or <4 WDI scores in each week. This will be presented by treatment group and overall treatment groups.

6.3. Appendix 3 Data Derivations Rule

6.3.1. Criteria for Potential Clinical Importance

6.3.1.1. Laboratory Values

Reference ranges for all laboratory parameters collected throughout the study are provided by the laboratory. A laboratory value that is outside the reference range is considered either high abnormal (value above the upper limit of the reference range) or low abnormal (value below the lower limit of the reference range). Note: a high abnormal or low abnormal laboratory value is not necessarily of clinical concern.

Laboratory Parameter	Units	Category	Clinical Concern Range	
			Low Flag (< x)	High Flag (>x)
Haematology				
Haematocrit	Ratio of 1		Decrease from baseline >0.075. and is below LLN	0.54
Haemoglobin	g/L		Decrease from baseline >25g/L	180
Lymphocytes	$x10^9$/ L		0.8	
Neutrophil Count	$x10^9$/ L		1.5	
Platelet Count	$x10^9$/ L		100	550
While Blood Cell Count	$x10^9$/ L		3	20
Coagulation				
INR	Ratio			> 1.5 (absolute value)
Clinical Chemistry				
Albumin	g/L		30	
		Δ from BL	>20% decrease from baseline	
Calcium	mmol/L		2	2.75
Creatinine	umol/L	Δ from BL		Increase from baseline >26.5
Glucose (fasting)	mmol/L		3	9
Potassium	mmol/L		3	5.5
Sodium	mmol/L		130	150
eGFR (CKD-EPI)	mL/min/ $1.73m^2$		30	
Liver Function				
ALT	IU/L			3x baseline
AST	IU/L			3x baseline
ALP	IU/L			2x baseline
Total Bilirubin	µmol/L			1.5x baseline
GGT	IU/L			2x baseline

6.3.1.2. Vital Signs

The following criteria will be used to flag potential clinical importance:

Vital Sign Parameter (Absolute)	Units	Clinical Concern Range	
		Lower	Upper
Systolic Blood Pressure	mmHg	< 85	> 160
Diastolic Blood Pressure	mmHg	< 45	> 100
Pulse Rate	bpm	< 40	> 110
QRS interval	Msec	< 75	> 110

6.3.2. Study Period

6.3.2.1. Classification Related to Study Period

Assessments and events will be classified according to the time of occurrence relative to study period.

Study Period	Treatment Details	Definition	Notes
Screening	None	Date < Date of Randomisation	
Part A	Part A	Date of Randomisation ≤ Date ≤ V9	Date of Randomisation is collected as 'Date of Treatment Assignment' on Treatment Assignment CRF page.
Part B	Part B	V9 < Date ≤ Date of Study Completion (or Discontinuation)	Date of Study Completion or Discontinuation as collected on Study Conclusion CRF page.

Note: Visit 9 dates will be provided via the SDTM supplementary SV domain and refer to the exact date of the participants visit.

Note: For visit-based assessment, the nominal visit name may be used to assign the treatment period (e.g., Visit 3 to Visit 9 to Part A, and Visit 10 and 11 to Part B).

6.3.2.2. Classification Related to Study Intervention State

Assessments and events will also be classified according to the time of occurrence relative to the study intervention period.

Treatment State	Definition
Pre-Treatment	Date < Date of start of study intervention
On-Treatment (Intervention emergent)	Date of start of study intervention ≤ Date ≤ Study intervention stop date (or Date of last dose of study intervention if lost to follow-up at the end of the study) If time of assessment or study intervention is not collected, the following assessment on the first dose date will be assumed to be taken prior to the first dose and therefore considered pre-intervention: ECG, Lab, and vital signs
Post-Treatment	Date of last dose of study intervention < Date

The definitions of treatment states for AEs are provided below:

Treatment State	Definition
Pre-Treatment	AE start date < Date of start of study intervention
On-Treatment Part A	For participants that switched from Part A to Part B treatment without any gaps in between the switch • Part A Study Treatment Start Date ≤ AE Start Date ≤ Part A Study Treatment Stop Date. For participants that switched from Part A to Part B treatment with a gap of less than or equal to two days in between the switch • Part A Study Treatment Start Date ≤ AE Start Date < Part B Study Treatment Start Date. For participants that switched from Part A to Part B treatment with a gap of more than two days in between the switch (or entered Part B but never started Part B treatment) • Part A Study Treatment Start Date ≤ AE Start Date ≤ Part A Study Treatment Stop Date + 2 days. For participants that withdraw from the study or discontinue treatment during Part A of the study • Part A Study Treatment Start Date ≤ AE Start Date ≤ Part A Study Treatment Stop Date + 2 days.
On-Treatment Part B	Date of start of study intervention Part B ≤ AE start date ≤ Part B Study Treatment Stop Date (or Date of last dose of study intervention Part B if lost to follow-up at the end of the study) + 2 days
Post-Treatment Part A	For participants that switched from Part A to Part B treatment with a gap of more than two days in between the switch • Part A Study Treatment Stop Date + 2 days < AE Start Date < Date of first dose of study intervention Part B For participants that withdraw from the study or discontinue treatment during Part A of the study • Part A Study Treatment Stop Date + 2 days < AE Start Date < Withdrawal date if available (otherwise, follow-up visit date or last contact date [whichever is later])
Post-Treatment Part B	Part B Study Treatment Stop Date (or Date of last dose of study intervention Part B if lost to follow-up at the end of the study) + 2 days < AE start date
Onset Time Since 1st Dose	If Date of start of study intervention ≤ AE start date: AE Start Date – Date of Start of Study Intervention +1 Missing otherwise.
Duration (Days)	AE Resolution Date – AE Onset Date + 1

6.3.2.3. Classification of Concomitant Medication

Medication State	Definition
Prior	Medication start date < Date of start of study intervention
Concomitant Part A	Medication start date ≤ Date of last dose of Part A study intervention, and Medication end date > Date of start of Part A study intervention. A medication could potentially be both prior and concomitant.
Concomitant Part B	Medication start date ≤ Date of last dose of Part B study intervention, and Medication end date > Date of start of Part B study intervention. A medication could potentially be both prior and concomitant.
Notes: If the medication start date is missing but it is flagged as ongoing, then the medication will be assumed to be ongoing throughout the study.	

6.3.3. Study Intervention Emergent Flag for Adverse Events

Flag	Definition
Study Intervention Emergent Part A	For participants that switched from Part A to Part B treatment without any gaps in between the switch • Part A Study Treatment Start Date ≤ AE Start Date ≤ Part A Study Treatment Stop Date. For participants that switched from Part A to Part B treatment with a gap of less than or equal to two days in between the switch ○ Part A Study Treatment Start Date ≤ AE Start Date < Part B Study Treatment Start Date. For participants that switched from Part A to Part B treatment with a gap of more than two days in between the switch (or entered Part B but never started Part B treatment) • Part A Study Treatment Start Date ≤ AE Start Date ≤ Part A Study Treatment Stop Date + 2 days. For participants that withdraw from the study or discontinue treatment during Part A of the study • Part A Study Treatment Start Date ≤ AE Start Date ≤ Part A Study Treatment Stop Date + 2 days.
Study Intervention Emergent Part B	• Part B Study Treatment Start Date ≤ AE Start Date ≤ Part B Study Treatment Stop Date + 2 days.

Notes:

- If the study treatment stop date is missing, then the AE will be considered to be intervention emergent.
- The Part B treatment start date is defined as the first day where the participant received only Part B treatment. For participants who switch from Part A treatment to Part B treatment on the same day, the Part B treatment start date will be the day following the switch.
- A gap is defined as a day where no treatment is taken.

6.3.4. Study Day and Reference Dates

The safety reference date is the study intervention start date and will be used to calculate study day for safety measures.

The efficacy reference date is the study intervention start date and will be used to calculate study day for efficacy measures and efficacy durations. The baseline characteristics will be calculated based on the date of randomization.

The study day is calculated as below with no day 0 defined:

- Assessment Date = Missing → Study Day = Missing
- Assessment Date < Reference Date → Study Day = Assessment Date – Ref Date
- Assessment Data ≥ Reference Date → Study Day = Assessment Date – Ref Date + 1

6.3.5. Analysis Window

For data summaries by visit, the nominal visit description will be used. Unscheduled and withdrawal visit data will be slotted into a target visit based on analysis window defined in the table below. If there are multiple assessments within the same window, a scheduled visit will be prioritized over un-scheduled visits. If all assessments within the same window are from unscheduled visits, the earliest one will be taken in the slotting. The reference date is the start of study intervention.

Analysis Set / Domain	Parameter (if applicable)	Target	Analysis Window		Analysis Timepoint
			Beginning Timepoint	Ending Timepoint	
Safety/ITT	All	Day 1	Day 1	Day 14	Visit 3
		Day 28	Day 15	Day 42	Visit 4
		Day 56	Day 43	Day 70	Visit 5
		Day 84	Day 71	Day 98	Visit 6
		Day 112	Day 99	Day 126	Visit 7
		Day 140	Day 127	Day 154	Visit 8
		Day 168	Day 155	Day 182	Visit 9
		Day 196	Day 183	Day 210	Visit 10
		Day 224	Day 211	Day 238	Visit 11

Note: Day 1 is the start of study intervention.

6.3.6. Study Population Data

6.3.6.1. Treatment Compliance

- Compliance by number of tablets

Compliance will be calculated based on the number of tablets dispensed and returned within Part A (dispensed at Visits 3, 4, and 6, and returned at Visits 4-9) and Part B (dispensed at Visit 9 and returned at Visits 10 - 11)

Part A Compliance =100 * (total number of tablets dispensed from Visit 3 to Visit 6 – total number of tablets returned from Visit 4 to Visit 9 / (expected number of tablets to be taken), where expected number of tablets to be taken = 2 tablets per day * (Visit 9 – Day 1 of study intervention).

Part B Compliance =100 * (total number of tablets dispensed at Visit 9 – total number of tablets returned from Visit 9 to Visit 11 / (expected number of tablets to be taken), where expected number of tablets to be taken = 2 tablets per day * (Visit 11 – Visit 9).

If a participant is lost to follow-up, compliance will be calculated up to the last Visit attended, assuming no tablets were taken from any bottles dispensed at this visit. These participants will be flagged, and a footnote added to any summaries of compliance to provide information on the number of participants where compliance has not been calculated over the entirety of the study part.

For example, if a participant does not return for Visit 6 and is lost to follow-up, their compliance during Part A will be calculated as: 100 * (total number of tablets dispensed from Visits 3 to 4 – total number of tablets returned from Visits 3 to 5) / (expected number of tablets to be taken), where expected number of tablets to be taken = 2 tablets per day * (Visit 4 – Visit 3), with a flag added to say that the participant drops out after Visit 4.

Similarly, if a participant withdraws from the study, compliance should be calculated up to the ED visit.

- Compliance by e-Dairy

The compliance will be calculated based on the e-Dairy questions.

For each day of Part A, from the evening of Day 1 of study intervention until the morning of Visit 9, the following will be calculated:

CEQam = 1 if 'Yes, I took both the morning and evening dose' is selected;

= 1 if 'I took the morning dose only' is selected;

= 0 if otherwise

CEQpm = 1 if 'Yes, I took both the morning and evening dose' is selected;

= 1 if 'I took the evening dose only' is selected;

= 0 if otherwise

Total Part A CEQam = Number of days that CEQam = 1 between Day 2 of study intervention and Visit 9.

Total Part A CEQpm = Number of days that CEQpm = 1 between Day 1 of study intervention and Visit 9 - 1.

Part A AM Compliance = 100 * (Total CEQam)/expected days, where expected days is Visit 9 – Day 1 of study intervention

Part A PM Compliance = 100 * (Total CEQpm)/ expected days, where expected days is Visit 9 – Day 1 of study intervention

Part A Compliance =mean (Part A AM compliance, Part A PM Compliance)

Similarly, for Part B,

Total Part B CEQam = Number of days that CEQam = 1 between Visit 9 + 1 and Visit 11.

Total Part B CEQpm = Number of days that CEQpm = 1 between Visit 9 + and Visit 11 - 1.

Part A AM Compliance = 100 * (Total CEQam)/expected days, where expected days is Visit 11 – Visit 9

Part B PM Compliance = 100 * (Total CEQpm)/ expected days, where expected days is Visit 11 – Visit 9

Part B Compliance =mean (Part B AM compliance, Part B PM Compliance)

If a participant is lost to follow-up, the compliance is calculated up to the date of the last visit, or the date of the last e-Dairy entry, whichever is later. If a participant withdraws from the study, the compliance is calculated up to the date of the ED visit. Similar as the

compliance by number of tablet, these participants will be flagged, and a footnote added to any summaries of compliance to provide information on the number of participants where compliance has not been calculated over the entirety of the study part.

6.3.6.2. Chronic Kidney Disease (CKD) Stage

The following CKD ranges will be used:

- Stage 1 (normal or high GFR): eGFR >=90 mL/min/1.73m^2
- Stage 2 (mild CKD): eGFR 60-89 mL/min/1.73m^2
- Stage 3a (moderate CKD): eGFR 45-59 mL/min/1.73m^2
- Stage 3b (moderate CKD): eGFR 30-44 mL/min/1.73m^2
- Stage 4 (severe CKD): eGFR 15-29 mL/min/1.73m^2
- Stage 5 (end stage CKD): eGFR <15 mL/min/1.73m^2

Note: eGFR forms an exclusion criterion for this study and therefore baseline values are expected to be ≥ 30.

6.3.7. Efficacy Data

6.3.7.1. Itch Score

6.3.7.1.1. Worst Daily Itch Score

The itch eDiary is to be completed twice each day to rate the worst bedtime itch (completed in the morning; AM score) and the worst daytime itch (completed during the evening; PM score) using 0-10 NRS score. The WDI score is the maximum (worst) of the AM and PM NRS scores.

If there is missing data within a day:

Missing bedtime Worst Itch score then	WDI = PM Worst Itch score
Missing daytime Worst Itch score then	WDI = AM Worst Itch score
Missing AM and PM Worst Itch score then	WDI = Missing

If there is more than one entry in the e-diary data for worst itch at a timepoint (AM or PM) of a given day, the worst (the highest) score from each duplicate record at that timepoint will be used in the derivation of WDI.

6.3.7.1.2. Weekly Itch score

The WIS is the average of the WDI scores in one week. The reference date of WIS during Part A (Week 1 – Week 24) is the start of study intervention (Day 1). The reference date of WIS during Part B (Week 25 – Week 32) is Visit 9.

The WIS at each week of Part A and B is calculated as the average of WDI from start date to end date shown below:

Analysis Time point	Start Date	End Date
Week 1	Day 2	Day 8
Week 2	Day 9	Day 15
Week 3	Day 16	Day 22
Week 4	Day 23	Day 29
Week 5	Day 30	Day 36
Same logic followed for Week 6 to Week 22		
Week 23	Day 156	Day 162
Week 24	Day 163	Day 169 or Visit 9 - 1, whichever is earlier
Week 25	Visit 9 + 1	Visit 9 + 7
Week 26	Visit 9 + 8	Visit 9 + 14
Same logic followed for Week 27 – Week 31		
Week 32	Visit 9 + 50	Visit 9 + 56 or Visit 11 – 1, whichever is earlier

Note: Day 1 is the start of study intervention in Part A.

The reference date of Part A baseline WIS is randomization day. The WIS at Part A baseline weeks are calculated as the average of WDI from start date to end date shown below:

Analysis Time point	Start Date	End Date
Part A Baseline Week 1	Day -28	Day -22
Part A Baseline Week 2	Day -21	Day -15
Part A Baseline Week 3	Day -14	Day -8
Part A Baseline Week 4	Day -7	Day -1

Note: Day -1 is the day immediately preceding randomization date.

At least 4 WDI scores are required for the calculation of WIS of that week. If more than 3 of the WDI scores are missing in any week, then the WIS of that week will be set to missing.

The WIS at Part A Baseline Week 4 is the baseline used for endpoints on change from baseline in WIS. Therefore, change from baseline in WIS at Week x = WIS at Week x – WIS at Part A Baseline Week 4.

6.3.7.1.3. Monthly Itch Score

The MIS is the worst WIS for that month (i.e., worst week score of the 4 weeks).

The MIS at a visit is the worst (maximum) WIS of the weeks defined above in 6.3.7.1.2, as shown below:

Analysis Time point	Labels for Summaries	Monthly Itch Score
Visit 3	Part A Baseline (V3)	Maximum (WIS at Part A Baseline Week 1 – 4)
Visit 4 / Week 4	Week 4 (V4)	Maximum (WIS at Week 1 – 4)
Visit 5 / Week 8	Week 8 (V5)	Maximum (WIS at Week 5 – 8)

Analysis Time point	Labels for Summaries	Monthly Itch Score
Visit 6 / Week 12	Week 12 (V6)	Maximum (WIS at Week 9 – 12)
Visit 7 / Week 16	Week 16 (V7)	Maximum (WIS at Week 13 – 16)
Visit 8 / Week 20	Week 20 (V8)	Maximum (WIS at Week 17 – 20)
Visit 9 / Week 24	Week 24 (V9)	Maximum (WIS at Week 21 – 24)
Visit 10 / Week 28	Week 28 (V10)	Maximum (WIS at Week 25 – 28)
Visit 11 / Week 32	Week 32 (V11)	Maximum (WIS at Week 29 – 32)

If one or more WIS are missing, then the worst (maximum) of the non-missing WIS is selected as the MIS. If no WIS are available for the 4-week periods defined in the table above, the MIS at that Visit will be set to missing.

The frequency and percentage of participants with 0, 1, 2, 3, or 4 weekly itch scores available at each month will be presented by treatment group.

The change from baseline in MIS at Visit x during Part A = MIS at Visit x – MIS at Part A Baseline (V3).

The baseline of MIS at Part B is Week 24 MIS. The change from baseline in MIS at Visit x during Part B = MIS at Visit x – MIS at Week 24 (V9)

6.3.7.2. Sleep Score

6.3.7.2.1. Daily Sleep Score

The sleep eDiary is to be completed every morning to score the severity of how itch interfered with sleep using an NRS from 0 to 10.

If there is more than one entry in the e-diary data for the sleep score on any given day, the worst score will be used as the DSS for that day.

6.3.7.2.2. Weekly Sleep Score

The WSS is the average of the DSS in one week. The reference date of WSS during Part A (Week 1 – Week 24) is the start of study intervention (Day 1).

The WSS at each week of part A is calculated as the average of DSS from start date to end date shown below:

Analysis Time point	Start Date	End Date
Week 1	Day 2	Day 8
Week 2	Day 9	Day 15
Week 3	Day 16	Day 22
Week 4	Day 23	Day 29
Week 5	Day 30	Day 36
Same logic followed for Week 6 to Week 22		
Week 23	Day 156	Day 162

Analysis Time point	Start Date	End Date
Week 24	Day 163	Day 169 or Visit 9 - 1, whichever is earlier

Note: Day 1 is the start of study intervention in Part A.

The reference date of Part A baseline WSS is randomization day. The WSS at Part A baseline weeks are calculated as the average of DSS from start date to end date shown below:

Analysis Time point	Start Date	End Date
Part A Baseline Week 1	Day -28	Day -22
Part A Baseline Week 2	Day -21	Day -15
Part A Baseline Week 3	Day -14	Day -8
Part A Baseline Week 4	Day -7	Day -1

Note: Day -1 is the day immediately preceding randomization date.

If more than 3 of the DSS are missing in any week then the WSS of that week will be set to missing.

The WSS at Part A Baseline Week 4 is the baseline used for endpoints on change from baseline in WSS. Therefore, the change from baseline in WSS at Week *x*= WSS at Week *x* – WSS at Part A Baseline Week 4.

6.3.7.2.3. Monthly Sleep Score

The MSS is the worst WSS for that month (i.e., worst week score of the 4 weeks). The MSS at a visit is the worst (maximum) WSS of the weeks defined above in 6.3.7.2.2, as shown below:

Analysis Time point	Labels for Summaries	Monthly Sleep Score
Visit 3	Part A Baseline (V3)	Maximum (WSS at Part A Baseline Week 1 – 4)
Visit 4 / Week 4	Week 4 (V4)	Maximum (WSS at Week 1 – 4)
Visit 5 / Week 8	Week 8 (V5)	Maximum (WSS at Week 5 – 8)
Visit 6 / Week 12	Week 12 (V6)	Maximum (WSS at Week 9 – 12)
Visit 7 / Week 16	Week 16 (V7)	Maximum (WSS at Week 13 – 16)
Visit 8 / Week 20	Week 20 (V8)	Maximum (WSS at Week 17 – 20)
Visit 9 / Week 24	Week 24 (V9)	Maximum (WSS at Week 21 – 24)

If one or more WSS scores are missing, then the worst (maximum) of the non-missing WSS scores is selected as the MSS. If no WSS scores are available for the 4-week periods defined in the table above, the MIS at that Visit will be set to missing.

The frequency and percentage of participants with 0, 1, 2, 3, or 4 weekly sleep scores available at each month will be presented by treatment group.

The change from baseline in MSS at Visit x= MSS at Visit x – MSS at Part A Baseline (V3).

6.3.7.3. Fatigue Score

6.3.7.3.1. Daily Fatigue Score

The fatigue eDiary is to be completed every evening to score the fatigue using an NRS from 0 to 10.

If there is more than one entry in the e-diary data for the fatigue score on any given day, the worst score will be used as the DFS for that day.

6.3.7.3.2. Weekly Fatigue Score

The WFS is the average of the DFS in one week. The reference date of WFS during Part A (Week 1 – Week 24) is the start of study intervention (Day 1).

The WFS at each week of part A is calculated as the average of DFS from start date to end date shown below:

Analysis Time point	Start Date	End Date
Week 1	Day 2	Day 8
Week 2	Day 9	Day 15
Week 3	Day 16	Day 22
Week 4	Day 23	Day 29
Week 5	Day 30	Day 36
Same logic followed for Week 6 to Week 22		
Week 23	Day 156	Day 162
Week 24	Day 163	Day 169 or Visit 9 - 1, whichever is earlier

Note: Day 1 is the start of study intervention in Part A.

The reference date of Part A baseline WFS is randomization day. The WFS at Part A baseline weeks are calculated as the average of DFS from start date to end date shown below:

Analysis Time point	Start Date	End Date
Part A Baseline Week 1	Day -28	Day -22
Part A Baseline Week 2	Day -21	Day -15
Part A Baseline Week 3	Day -14	Day -8
Part A Baseline Week 4	Day -7	Day -1

Note: Day -1 is the day immediately preceding randomization date.

If more than 3 of the DFS are missing in any week then the WFS of that week will be set to missing.

The WFS at Part A Baseline Week 4 is the baseline used for endpoints on change from baseline in WFS. Therefore, the change from baseline in WFS at Week x= WFS at Week x – WFS at Part A Baseline Week 4.

6.3.7.3.3. Monthly Fatigue Score

The MFS is the worst WFS for that month (i.e., worst week score of the 4 weeks). The MFS at a visit is the worst (maximum) WFS of the weeks defined above in 6.3.7.3.2, as shown below:

Analysis Time point	Labels for Summaries	Monthly Fatigue Score
Visit 3	Part A Baseline (V3)	Maximum (WFS at Part A Baseline Week 1 – 4)
Visit 4 / Week 4	Week 4 (V4)	Maximum (WFS at Week 1 – 4)
Visit 5 / Week 8	Week 8 (V5)	Maximum (WFS at Week 5 – 8)
Visit 6 / Week 12	Week 12 (V6)	Maximum (WFS at Week 9 – 12)
Visit 7 / Week 16	Week 16 (V7)	Maximum (WFS at Week 13 – 16)
Visit 8 / Week 20	Week 20 (V8)	Maximum (WFS at Week 17 – 20)
Visit 9 / Week 24	Week 24 (V9)	Maximum (WFS at Week 21 – 24)

If one or more WFS scores are missing, then the worst (maximum) of the non-missing WFS scores is selected as the MFS. If no WFS scores are available for the 4-week periods defined in the table above, the MFS at that Visit will be set to missing.

The change from baseline in MFS at Visit x= MFS at Visit x – MFS at Part A Baseline (V3).

6.3.7.4. Itch Responder

Improvement in MIS score is the reduction of the score (10 is the highest and worst severity, 0 lowest and least severe).

Responders at Visits 4 – 9 (Week 4, 8, 12, 16, 20, and 24) are evaluated by comparing the change from baseline in MIS at that visit against the responder threshold.

For responder threshold of achieving a ≥2-point reduction from Baseline in the Monthly Itch score,

Responder = 1, if change from baseline in MIS ≤ -2

= 0, if change from baseline in MIS > -2

= missing, if MIS is missing

For responder threshold of achieving a ≥3-point reduction from Baseline in the Monthly Itch score,

Responder = 1, if change from baseline in MIS ≤ -3

= 0, if change from baseline in MIS > -3

= missing, if MIS is missing

For responder threshold of achieving a ≥4-point reduction from Baseline in the Monthly Itch score,

Responder = 1, if change from baseline in MIS ≤ -4

= 0, if change from baseline in MIS > -4

= missing, if MIS is missing

The missing MIS values will be imputed before dichotomized to responder for the CMH analysis as detailed in Section 4.3.1

6.3.7.5. Itch Response Day

Itch response at a day is defined by comparing the change from baseline WDI score at a day to an itch response threshold. Baseline is defined as the WIS of 7 days prior to randomization.

For threshold of achieving a ≥2-point reduction from Baseline in WDI,

WDI response = 1, if change from baseline in WDI score ≤ -2

= 0, if change from baseline in WDI score > -2

= missing, if WDI is missing

For threshold of achieving a ≥3-point reduction from Baseline in WDI,

WDI response = 1, if change from baseline in WDI score ≤ -3

= 0, if change from baseline in WDI score > -3

= missing, if WDI is missing

For threshold of achieving a ≥4-point reduction from Baseline in WDI,

WDI response = 1, if change from baseline in WDI score ≤ -4

= 0, if change from baseline in WDI score > -4

= missing, if WDI is missing

The percentage of itch response days during part A is calculated, for each itch response threshold, as

(Number of response days / total number days from Day 2 to V9 -1 for which a WDI score is available) * 100

6.3.7.6. PBC-40

The answers to the individual PBC-40 questions should be scored as in http://www.uk-pbc.com/wp-content/uploads/2015/12/coded-PBC-40.pdf (summarized below).

Question 1 is scored as:

5 – Never
4 – Rarely
3 – Sometime
2 – Most of the Time
1 – Always

Question 2 to 27 (inclusive) are scored as:

1 – Never
2 – Rarely
3 – Sometime
4 – Most of the Time
5 – Always

Questions 28 to 33 (inclusive) are scored as:

1 – Not at all
2 – A little
3 – Somewhat
4 – Quite a bit
5 – Very much

Questions 34-39 (inclusive) are scored as:

5 – Strongly Agree
4 – Agree
3 – Neither Agree nor Disagree
2 – Disagree
1 – Strongly Disagree

Question 40 is scored as:

1 – Strongly Agree
2 – Agree
3 – Neither Agree nor Disagree
4 – Disagree
5 – Strongly Disagree

For all questions, an answer of Does/Did not apply will be scored 0.

For each visit the PBC-40 questionnaire should provide scores for 6 domains:

PBC-Symptoms = sum of scores of questions 1 to 7

PBC-Itch = sum of scores of questions 8 to 10

PBC Fatigue = sum of scores of questions 11 to 21

PBC-Cognitive = sum of scores of questions 22 to 27

PBC-Emotional = sum of scores of questions 28, 30 and 33

PBC-Social = sum of scores of questions 29, 31 and 32, and 34 to 40

For each individual PBC domain (symptoms, itch, emotional, social, fatigue and cognitive), if more than 50% of the corresponding question scores for a participant are missing then the total score for that domain will be marked as missing. If 50% or more of the corresponding question scores are complete, the median of these responses will be imputed for the missing values within the domain before calculating the domain score.

6.3.7.7. Patient Global Impression of Severity

Derived PGI-S scores will be on a scale of 1-5 from the PGI-S question:

1 = Absent
2 = Mild
3 = Moderate
4 = Severe
5 = Very Severe

6.3.7.8. Patient Global Impression of Change

Derived PGI-C score will be on a scale of 1-7 from the PGI-C question:

1 = Very Much Improved
2 = Moderately Improved
3 = Minimally Improved
4 = No Change
5 = Minimally Worse
6 = Moderately Worse
7 = Very Much Worse

6.3.7.9. Epworth Sleepiness Scale

The ESS consists of 8 questions using scales ranging from 0 to 3, where scale of 0 means "would never doze" and scale of 3 means "high chance of dozing". The total ESS score is the sum of the scores of all questions and ranges from 0 to 24.

If there is more than one score made for question, the mean score will be used.

If there are decimal scores entered for questions, the scores will be taken at face value and added to compose the total score. If the total score includes a decimal, it will be rounded to form whole number.

If one or more responses are missing, the total score is missing.

From the author, the total ESS scores can be interpreted as follows:

- 0-5: Lower Normal Daytime Sleepiness
- 6-10: Higher Normal Daytime Sleepiness
- 11-12: Mild Excessive Daytime Sleepiness
- 13-15: Moderate Excessive Daytime Sleepiness
- 16-24: Severe Excessive Daytime Sleepiness

6.3.7.10. Beck Depression Inventory –II

The BDI-II consists of 21 questions, with possible responses being scored from 0 (least severe response) to 3 (most severe response). The scores for all 21 questions are then added together to obtain an 'observed total score' on a scale of 0 to 63.

If there are no missing responses then

Total Score = Observed Total Score

If there are 1 or 2 missing responses the Total Score can be adjusted to account for the missing responses by:

Total Score = (21 / number of non-missing responses) * Observed Total Score

If 3 or more responses are missing then:

Total Score = missing.

IMPORTANT NOTE: For questions 16 and 18, the collected code values must be adjusted as shown below to compute the item score:

Question 16:

Collection Code	Label	Item Score
0	0 - I have not experienced any change in my sleeping pattern	0
1	1a - I sleep somewhat more than usual	1
2	1b - I sleep somewhat less than usual	1
3	2a - I sleep a lot more than usual	2
4	2b - I sleep a lot less than usual	2
5	3a - I sleep most of the day	3
6	3b - I wake up 1-2 hours early and can't get back to sleep	3

Question 18:

Collection Code	Label	Item Score
0	0 - I have not experienced any change in my appetite	0
1	1a - My appetite is somewhat less than usual	1
2	1b - My appetite is somewhat greater than usual	1
3	2a - My appetite is much less than before	2
4	2b - My appetite is much greater than usual	2
5	3a - I have no appetite at all	3
6	3b - I crave food all the time	3

6.3.7.11. Discard of Data based on Hypothetical Estimand Strategy

- Hypothetical estimand strategy for COVID-19 pandemic related intercurrent events

For the primary and key secondary endpoints where hypothetical strategy is applied to COVID-19 pandemic related intercurrent events (see Section 1.1 for more details), WDI scores or DSS impacted by study treatment interruption due to COVID-19 pandemic will be discarded and treated as missing. The corresponding WIS or WSS will then be constructed, followed by the derivation of MIS or MSS. See Section 4.2.1 for more details.

- Hypothetical estimand strategy for any intercurrent events

For the supplementary estimand of the primary endpoint, hypothetical strategy will be used for any intercurrent events, including treatment discontinuation, disruptions in treatment or treatment delays, changes in itch therapy or use of rescue medication. WDI scores impacted by these intercurrent events will be discarded and treated as missing. The corresponding WIS will then be constructed, followed by the derivation of MIS.

6.3.8. Safety Data

6.3.8.1. Exposure calculation

Number of days of exposure to study drug during part A and B will be calculated based on the formula below regardless of temporary treatment interruption:

Part A Duration of Exposure in Days = Part A Treatment Stop Date – Part A Treatment Start Date + 1

Part B Duration of Exposure in Days = Part B Treatment Stop Date – Part B Treatment Start Date + 1

Participants who were randomized but did not report a treatment start date will be categorized as having zero days of exposure.

For participants with treatment stop date missing, if treatment stop occurs during Part A, the stop date will be imputed as Visit 9 date or early discontinuation/withdrawal (ED) date, whichever is earlier.

If there is no evidence that the treatment stops during Part A, the stop date will be imputed as Visit 11 date or ED date, whichever is earlier.

Note that the Part B treatment start date is defined as the first day where the participant received only Part B treatment. For participants who switch from Part A treatment to Part B treatment on the same day, the Part B treatment start date will be the day following the switch.

The cumulative dose will be based on the formula:

Cumulative Dose = Sum of (Duration of Exposure x Total Daily Dose)

6.3.8.2. Exposure adjusted incidence rate

When cumulative incidence rate is reported for AE analyses, exposure adjusted incidence rate will also be reported. Exposure adjusted incidence rate will be calculated based on the formula below:

$$Exposure\ adjusted\ incidence\ rate = \frac{Number\ of\ patients\ with\ the\ event}{Total\ exposure\ time\ in\ year} \times 100$$

6.3.8.2.1. Exposure adjusted incidence rate for the Main Safety Estimand

For the main safety estimand, a while-on-treatment strategy will be used to address permanent treatment discontinuation. Exposure will be calculated as follows:

For AEs reported in Part A, Total exposure time in year = sum of participant total exposure duration in days during Part A for all safety population (across all participants)/365.25, where participant total exposure duration in days is calculated as follows:

- if they DO experience the AE, then start date of first AE – Part A treatment start date + 1.
- if they DO NOT experience the AE and DO NOT Withdraw from the study during Part A, and they switched from Part A to Part B treatment without any gaps in between the switch, then **Part A Duration of Exposure in Days** (See Section 6.3.8.1)
- if they DO NOT experience the AE and DO NOT Withdraw from the study during Part A, and they switched from Part A to Part B treatment with a gap of less than or equal to two days in between the switch, then ([Part B treatment start date – 1 Day] – Part A treatment start date) + 1
- if they DO NOT experience the AE and DO NOT Withdraw from the study during Part A, and they switched from Part A to Part B treatment with a gap of more than two days in between the switch (or entered Part B but never started Part B treatment), then **Part A Duration of Exposure in Days** (See Section 6.3.8.1) + 2 Days.

- If they DO NOT experience the AE and DO Withdraw from the study during Part A, then **Part A Duration of Exposure in Days** (See Section 6.3.8.1) + 2 Days.

Similarly, for AEs reported in Part B, Total exposure time in year = sum of participant total exposure duration in days during Part B for all safety population (across all participants)/365.25, where participant total exposure duration in days is calculated as follows:

- if they DO experience the AE, then start date of first AE – Part B treatment start date + 1.
- if they DO NOT experience the AE, then **Part B Duration of Exposure in Days** (See Section 6.3.8.1) + 2 Days.

Note that if a participant experiences multiple AEs in the same System Organ Class and Preferred Term, exposure adjusted incidence rates are only calculated for the first instance of the AE.

6.3.8.2.2. Exposure adjusted incidence rate for Supplementary Safety Estimand

For the supplementary safety estimand, a treatment policy strategy (or 'while-on-study') will be used to address permanent treatment discontinuation (i.e., including AEs after permanent treatment discontinuation). Exposure is defined from first date of treatment to the end of participation in the study, regardless of the end of exposure to treatment. Exposure will be calculated as follows:

For AEs reported in Part A, Total exposure time in year = sum of participant total exposure duration in days during Part A for all safety population (across all participants)/365.25, where participant total exposure duration in days is calculated as follows:

- If they DO experience the AE, then start date of first AE – Part A treatment start date + 1.
- if they DO NOT experience the AE and Withdraw from the study >2 days after Part A treatment stop date (but before end of visit 9), then withdrawal date – Part A treatment start date + 1
- If they DO NOT experience the AE and Withdraw from the study ≤ 2 days after Part A treatment stop date (but before end of visit 9), then (Part A treatment stop date + 2 – Part A treatment start date) + 1
- if they DO NOT experience the AE and DO NOT Withdraw from the study and DO NOT Discontinue treatment prior to starting Part B, then ([Part B treatment start date – 1 Day] – Part A treatment start date) + 1
- if they DO NOT experience the AE and DO NOT Withdraw from the study and DO Discontinue treatment before starting Part B (i.e., they never start Part B treatment), then Follow-up visit date or last contact (whichever is later) – Part A treatment start date + 1

Similarly, for AEs reported in Part B, Total exposure time in year = sum of participant total exposure duration in days during Part B for all safety population (across all participants)/365.25, where participant total exposure duration in days is calculated as follows:

- if they DO experience the AE, then start date of first AE – Part B treatment start date + 1
- if they DO NOT experience the AE and Withdraw from the study >2 days after Part B treatment stop date, then withdrawal date – Part B treatment start date + 1
- If they DO NOT experience the AE and Withdraw from the study ≤ 2 days after Part B treatment stop date, then (Part B treatment stop date + 2 – Part B treatment start date) + 1
- if they DO NOT experience the AE and DO NOT Withdraw from the study, then Follow-up visit date or Visit 11 Completion date (whichever is later) – Part B treatment start date + 1

Note that if a participant experiences multiple AEs in the same System Organ Class and Preferred Term, exposure adjusted incidence rates are only calculated for the first instance of the AE.

6.3.8.3. Gastrointestinal Symptoms Rating Scale

The responses to each question in the GSRS range from No discomfort at all (1) to Very severe discomfort (7).

For each week, the GSRS questionnaire should provide scores for 5 domains:

Diarrhoea Syndrome = average of scores for questions 11, 12 and 14

Indigestion Syndrome = average of scores for questions 6, 7, 8 and 9

Constipation Syndrome = average of scores for questions 10, 13 and 15

Abdominal pain syndrome = average of scores for questions 1, 4 and 5

Reflux Syndrome = average of scores for questions 2 and 3

Total = average scores for questions 1-15

For each individual GSRS domain (Diarrhoea, Indigestion, Constipation, Abdominal Pain, Reflux) if more than 50% of the corresponding question scores for a participant are missing then the total score for that domain will be marked as missing. If 50% or more of the corresponding question scores are complete the mean of these responses will be imputed for the missing values within the domain before calculating the domain score.

If there is more than one entry for the GSRS questionnaires on any given date or timepoint the worst score will be used for that date or timepoint as conservative approach.

As per the schedule of assessment, the GSRS will be completed at Visit 3 (prior to first dose), then every 7 days after the first dose of Part A treatment:

Analysis Time point	Start Date	End Date
Baseline	Visit 3	Visit 3
Week 1	Day 2	Day 8
Week 2	Day 9	Day 15
Week 3	Day 16	Day 22
Week 4	Day 23	Day 29
Week 5	Day 30	Day 36
Same logic followed for Week 6 to Week 22		
Week 23	Day 156	Day 162
Week 24	Day 163	Day 169 or Visit 9 - 1, whichever is earlier
Week 25	Visit 9 + 1	Visit 9 + 7
Week 26	Visit 9 + 8	Visit 9 + 14
Same logic followed for Week 27 – Week 31		
Week 32	Visit 9 + 50	Visit 9 + 56 or Visit 11 – 1, whichever is earlier

Note: Day 1 is the start of study intervention in Part A.

6.3.8.4. ECG Parameters

6.3.8.4.1. RR Interval

If RR interval (msec) is not provided directly, then RR can be derived based on QT and QTcF (or QTcB if available instead).

If QTcB is machine read & QTcF is not provided, then :

$$RR = \left[\left(\frac{QT}{QTcB}\right)^2\right] * 1000$$

Otherwise if QTcF is machine read, then:

$$RR = \left[\left(\frac{QT}{QTcF}\right)^3\right] * 1000$$

If both QTcB and QTcF are available QTcF should be used.

If ECGs are manually read, the RR value preceding the measurement QT interval will not be derived.

6.3.8.4.2. Corrected QT Intervals

When not entered directly in the eCRF, corrected QT intervals by Bazett's (QTcB) and Fredericia's (QTcF) formulas will be calculated, in msec, depending on the availability of other measurements.

If RR interval (msec) is provided then missing QTcB and/or QTcF will be derived as:

$$QTcB = \frac{QT}{\sqrt{\frac{RR}{1000}}}$$

$$QTcF = \frac{QT}{\sqrt[3]{\frac{RR}{1000}}}$$

6.3.8.5. Adverse Events of Special Interest

6.3.8.5.1. Diarrhea reported as an AE

Diarrhea AESI will be all study intervention emergent AEs with High Level Term (HLT) of "Diarrhoea (excl infective)" that have been marked as AESIs by investigators, reported via the Adverse Event eCRF.

6.3.8.5.2. Elevated ALT reported as a Liver Event (meets Liver Monitoring or Stopping criteria)

Elevated ALT reported as a Liver Event (meets Liver Monitoring or Stopping criteria) AESI will be all study intervention emergent AEs with System Organ Class (SOC) of "Hepatobiliary disorders" OR "Investigations" that have been marked as AESIs by investigators, as reported via the Adverse Event eCRF.

The study team will monitor liver event reporting to ensure that elevated ALT liver events are correctly reported as an AE and marked as an AESI on the AE eCRF.

6.3.8.6. Liver related SMQs

Liver related AEs will be summarized by standardized MedDRA Queries (SMQs). The following SMQs will be reported:

	SMQ
1	Cholestasis and jaundice of hepatic origin (SMQ).
2	Drug related hepatic disorders-severe events only (SMQ)
2a	Hepatic failure, fibrosis and cirrhosis and other liver damage related conditions (SMQ)
2b	Hepatitis, non-infectious (SMQ)
3	Liver related investigations, signs and symptoms (SMQ)
4	Liver related coagulation and bleeding disturbance (SMQ)

As per GSK SOP the MedDRA dictionary is reviewed and updated every 6 months. As a result, the list of terms that will be used to identify these SMQs will be updated on a 6 monthly basis following the up versioning of the dictionary. The final list of terms to be used will be confirmed at end of study.

6.3.8.7. Rescue Medication

Rescue medication is permitted only for participants who are not receiving background cholestatic pruritus treatment when entering the study. Participants using protocol defined rescue medication will be defined as follows:

Initiation of cholestyramine at or after Week 12, where;

- their Monthly Itch Score at or after Week 12 (Visit 6) has increased by at least 2 points compared to baseline AND their Monthly Itch Score is ≥ 7

OR

- their Baseline Monthly Itch Score > 8, AND their Monthly Itch Score at or after Week 12 (Visit 6) is 10

6.3.9. Laboratory Parameters

If a laboratory value which is expected to have a numeric value for summary purposes, has a non-detectable level reported in the database, where the numeric value is missing, but typically a character value starting with '<x' or '>x' (or indicated as less than x or greater than x in the comment field) is present, the number of decimal places in the observed values will be used to determine how much to add or subtract in order to impute the corresponding numeric value.

Example 1: 2 Decimal Places = '< x ' becomes x – 0.001 e.g. <0.05 becomes 0.049

Example 2: 1 Decimal Place = '< x' becomes x – 0.01 e.g. <0.2 becomes 0.19

Example 3: 0 Decimal Places = '> x' becomes x + 0.1 e.g. >99 becomes 99.1

6.3.10. Pharmacokinetic Concentrations

If linerixibat PK concentration is reported as missing in the linerixibat treatment group, then this will be reported as missing, if a concentration at V5, V6, V7, V8 or V9 is reported as non-quantifiable (NQ), then this will be imputed as ½ lower limit of quantification (LLQ) when the sample time is after dosing time, for inclusion in plots and PK concentrations summary tables.

6.3.11. Analysis Variables / Covariates Derivation

Variable/Covariate	Derivation
Baseline cholestyramine use (3 Levels: Previous cholestyramine use (no concomitant use at baseline); Baseline concomitant cholestyramine; Baseline cholestyramine naive)	**Previous cholestyramine use** (no concomitant use at baseline): The Previous PBC/Pruritus Treatments eCRF will be used. Where "Cholestyramine" is selected from the Bile Acid Binding Resins medications, and "Yes" is selected for the "Subject Stopped Treatment" field. The Concomitant Medication/Therapy eCRF data will also be checked to ensure that the medication end date is prior to randomisation. **Baseline concomitant cholestyramine:** The Previous PBC/Pruritus Treatments eCRF will be used. Where "Cholestyramine" is selected from the Bile Acid Binding Resins medications, and "No" is selected for the "Subject Stopped Treatment" field. The Concomitant Medication/Therapy eCRF data will also be checked. **Baseline cholestyramine naive:** Participants will be considered baseline cholestyramine naive if they do not meet either of the above (i.e., no previous Cholestyramine use and no baseline Cholestyramine use). The Concomitant Medication/Therapy eCRF data will also be checked to ensure there is no reported cholestyramine use.
Baseline concomitant cholestatic pruritus therapy (3 Levels: Regimen contains BABR; Regimen that does not contain BABR; No defined anti-cholestatic pruritus treatment)	**Regimen contains BABR:** The Previous PBC/Pruritus Treatments eCRF will be used. Where "Cholestyramine", "Colesevelam", "Colestipol" or "Colestilan/Colestimide" is selected from the Bile Acid Binding Resins medications, and "No" is selected for the "Subject Stopped Treatment" field. The Concomitant Medication/Therapy eCRF data will also be checked for any medications identified as: • **Bile acid binding resins** **Regimen that does not contain BABR;** The Previous PBC/Pruritus Treatments eCRF will be used. Where any medication that does not

Variable/Covariate	Derivation
	contain BABR is selected, and "No" is selected for the "Subject Stopped Treatment" field. The Concomitant Medication/Therapy eCRF data will also be checked for any medications identified as: • **Antihistamines:** Only when the indication is for Itch, as described in Section 4.6.1. • **Fibrates** • **Rifampicin/rifampin** • **Selective serotonin reuptake inhibitors** (e.g., sertraline, citalopram, fluoxetine, paroxetine, etc.): • **Pregabalin** • **Gabapentin** • **Opiate antagonists** (e.g., naltrexone, naloxone, nalfurafine, nalmefene) **No defined anti-cholestatic pruritus treatment:** All remaining participants where neither of the above are identified.
Baseline concomitant cholestatic pruritus therapy (Yes vs. No)	This will be derived using the "Baseline concomitant cholestatic pruritus therapy" variable above where: **Yes:** When Baseline concomitant cholestatic pruritus therapy = "Regimen contains BABR" OR "Regimen that doesn't contain BABR" **No:** When Baseline concomitant cholestatic pruritus therapy = "No defined anti-cholestatic pruritus treatment"
Baseline concomitant BABR (Yes vs. No)	This will be derived using the "Baseline concomitant cholestatic pruritus therapy" variable above where: **Yes:** When Baseline concomitant cholestatic pruritus therapy = "Regimen contains BABR"

Variable/Covariate	Derivation
	No: When Baseline concomitant cholestatic pruritus therapy = “No defined anti-cholestatic pruritus treatment” OR “Regimen that doesn’t contain BABR”
Baseline concomitant Fibrates (Yes vs. No)	**Yes:** The Previous PBC/Pruritus Treatments eCRF will be used. Where "Fenofibrate", "Bezafibrate" or "Other" is selected from the "Fibrates" medications, and “No” is selected for the “Subject Stopped Treatment” field. The Concomitant Medication/Therapy eCRF data will also be checked for any medications identified as: • **Fibrates** **No:** All remaining participants where no Fibrate use at baseline is identified.
Operating Model (3 Level: Metasite, Flexible, Brick and Mortar)	**Metasite:** There is only one metasite planned for this study (site number= 251379), therefore any participants where ‘SITEID=251379’ will be classed as ‘Metasite’. **Flexible:** Participants where SITEID ne 251379 AND have at least 1 study visit conducted remotely during Part A will be classed as ‘Flexible’. **Brick and mortar:** All remaining participants, where SITEID ne 251379 AND who have conducted all visits onsite (i.e, no remote visits during Part A) will be considered ‘Brick and mortar’. The number of Visits conducted remotely will be captured via the Visit Conduct eCRF at each visit using the question “visit conducted remotely” (Yes/No). We will use the response to this question to flag if any remote visits have been conducted.

6.3.12. Intercurrent Events

Intercurrent Event	Derivation
Permanent treatment discontinuation, disruptions in treatment or treatment delays unrelated to the COVID-19 pandemic	**Permanent discontinuation unrelated to COVID-19 pandemic:** • Permanent treatment discontinuation (i.e., treatment withdrawal or Study withdrawal) will be identified via the Study Conclusion eCRF and Study Treatment Discontinuation eCRF where they have also indicated that the discontinuation was not due to the COVID-19 pandemic – via the question "*Participant Discontinued From the Study/treatment Due to COVID-19 Pandemic*". • The Intercurrent event date will be the date of discontinuation. **Disruptions or delays in treatment unrelated to COVID-19 pandemic:** • Identified using the Study Treatment - Continual Dosing eCRF. Missed treatment start and end dates will be provided. • A disruption will be considered if at least 1 day (both doses missed) of treatment is missed. • When 1 or more days of treatment are missed, the start and end date of the intercurrent event will correspond to the start and end dates of the missed treatment, respectively.
Permanent treatment discontinuation, disruptions in treatment, or treatment delays related to the COVID-19 pandemic	**Permanent discontinuation related to COVID-19 pandemic:** • Permanent treatment discontinuation (i.e., treatment withdrawal or Study withdrawal) will be identified via the Study Conclusion eCRF and Study Treatment Discontinuation CRF where they have also indicated that the discontinuation was due to the COVID-19 pandemic – via the question "*Participant Discontinued From the Study/treatment Due to COVID-19 Pandemic*". • The Intercurrent event date will be the date of discontinuation.

Intercurrent Event	Derivation
	Disruptions or delays in treatment related to COVID-19 pandemic: • Identified using the Study Treatment Supply Interruption Due To COVID-19 Pandemic eCRF. Interruption start and end dates will be provided. • A disruption will be considered if at least 1 day (both doses missed) of treatment is missed. • When 1 or more days of treatment are missed, the start and end date of the intercurrent event will correspond to the start and end dates of the missed treatment, respectively.
Change in background itch therapy or use of rescue medication	• Concomitant Medication/Therapy eCRF will be used to identify Changes to background itch therapy or use of rescue since randomisation. • Starting/stopping or changing the dose of any itch medications will be considered as a change in background itch therapy. • The start date of the first change in background itch medication will be considered as the date of the intercurrent event (i.e., if there are multiple changes in background medication, then the first change will be considered, and Itch scores following this will be considered as being impacted). • If there is a partial start/end date, then dates will be imputed as per Section 6.3.14.

6.3.13. Multiple measurements at One Analysis Time Point

When triplicate ECG assessments are taken, mean of the measurement will be calculated first and summary statistics will be based on the calculated mean. This will apply to both baseline and post baseline assessments.

For lab tests on a study day, if more than one assessment is taken on the same day, the test from a central lab will be taken over the test from a local lab. If multiple assessments are taken from the same type of lab test, the worst case will be used.

Participants having both High and Low values versus Normal Ranges at any post-baseline visits for safety parameters will be counted in both the High and Low categories of "Any visit post-baseline" row of related summary tables. This will also be applicable to relevant Potential Clinical Importance summary tables.

If there is more than one entry in the e-diary data for itch, sleep, or fatigue on any timepoint (AM or PM) of a day, then the worst score for that day or timepoint will be used as described in Section 6.3.7.

If there is more than one entry for the GSRS questionnaire on any given date or timepoint, the worst score will be used for that date or timepoint as conservative approach.

6.3.14. Handling of Partial Dates

<table>
<tr><th>Element</th><th>Reporting Detail</th></tr>
<tr><td>General</td><td>

Partial dates will be displayed as captured in participant listing displays.
However, where necessary, display macros may impute dates as temporary variables for sorting data in listings only. In addition, partial dates may be imputed for ‘slotting’ data to study phases or for specific analysis purposes as outlined below.
For Adverse Events, imputed partial dates will not be used to derive study day, time to onset or duration (e.g., time to onset or duration of adverse events), or elapsed time variables (e.g., time since diagnosis). In addition, imputed dates are not used for deriving the last contact date in overall survival analysis dataset.
For Medical History, partial dates will be imputed for the calculation of disease duration.

</td></tr>
<tr><td>Adverse Events</td><td>

Partial dates for AE recorded in the CRF will be imputed using the following conventions:

<table>
<tr><td>Missing start day</td><td>If study intervention start date is missing (i.e. participant did not start study intervention), then set start date = 1st of month.

Else if study intervention start date is not missing:
If month and year of start date = month and year of study intervention start date, then
If stop date contains a full date and stop date is earlier than study intervention start date, then set start date= 1st of month.
Else set start date = study intervention start date.
Else set start date = 1st of month.</td></tr>
<tr><td>Missing start day and month</td><td>If study intervention start date is missing (i.e. participant did not start study intervention), then set start date = January 1.</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Element</th><th colspan="2">Reporting Detail</th></tr>
<tr><td></td><td></td><td>Else if study intervention start date is not missing:
o If year of start date = year of study intervention start date, then
▪ If stop date contains a full date and stop date is earlier than study intervention start date, then set start date = January 1.
▪ Else set start date = study intervention start date.
Else set start date = January 1.</td></tr>
<tr><td></td><td>Missing end day</td><td>A '28/29/30/31' will be used for the day (dependent on the month and year).</td></tr>
<tr><td></td><td>Missing end day and month</td><td>No Imputation</td></tr>
<tr><td></td><td>Completely missing start/end date</td><td>No imputation</td></tr>
<tr><td>Concomitant Medications/Medical History</td><td colspan="2">o Partial dates for any concomitant medications recorded in the CRF will be imputed using the following convention:
o For the calculation of Duration of PBC and Duration of Pruritus, partial dates will be imputed using the below convention:

<table><tr><td>Missing start day</td><td>If study intervention start date is missing (i.e. participant did not start study intervention), then set start date = 1st of month.
Else if study intervention start date is not missing:
o If month and year of start date = month and year of study intervention start date, then
▪ If stop date contains a full date and stop date is earlier than study intervention start date, then set start date= 1st of month.
▪ Else set start date = study intervention start date.
Else set start date = 1st of month.</td></tr><tr><td>Missing start day and month</td><td>If study intervention start date is missing (i.e. participant did not start study intervention), then set start date = January 1.
Else if study intervention start date is not missing:</td></tr></table></td></tr>
</table>

Element	Reporting Detail	
		○ If year of start date = year of study intervention start date, then • If stop date contains a full date and stop date is earlier than study intervention start date, then set start date = January 1. • Else set start date = study. intervention start date. Else set start date = January 1.
	Missing end day	A '28/29/30/31' will be used for the day (dependent on the month and year).
	Missing end day and month	A '31' will be used for the day and 'Dec' will be used for the month.
	Completely missing start/end date	No imputation

6.3.15. Trademarks

Trademarks of the GlaxoSmithKline Group of Companies	Trademarks not owned by the GlaxoSmithKline Group of Companies
	Fibroscan
	SAS
NONE	Enhanced Liver Fibrosis

7. REFERENCES

Apostolaros M, Babaian D, Corneli A, Forrest A, Hamre G, Hewett J, et al. Legal, regulatory, and practical issues to consider when adopting decentralized clinical trials: recommendations from the Clinical Trials Transformation Initiative. Therapeutic Innovation & Regulatory Science. 2020; 54:779-787.

GlaxoSmithKline Document Number TMF-16125967: Study 212620 IDMC Charter V2. 05 May 2023

GlaxoSmithKline Document Number TMF-16740250: Study 212620 protocol amendment 04. 20 Nov 2023

Wang Y, Tu W, Koh W, et al. Bayesian Hierarchical Models for Subgroup Analysis. Pharm Stat. Published online July 15, 2024. doi:10.1002/pst.2424