| | | |
|---|---|---|
| **Division** | : | Worldwide Development |
| **Information Type** | : | Reporting and Analysis Plan (RAP) |

| | | |
|---|---|---|
| **Title** | : | Reporting and Analysis Plan for Reporting and Analysis Plan for Study 201000: A randomised, double-blind, multi-dose, placebo-controlled study to evaluate the efficacy, safety and tolerability of GSK2330672 administration for the treatment of pruritus in participants with primary biliary cholangitis. |
| **Compound Number** | : | GSK2330672 |
| **Effective Date** | : | Refer to Document date |

**Description:**

- The purpose of this RAP is to describe the planned analyses and output to be included in the Clinical Study Report for Protocol 201000
- This RAP will be provided to the study team members to convey the content of the Statistical Analysis Complete (SAC) deliverable.
- The interim analysis was outlined in a separate Interim Analysis Plan document
- This RAP will be provided to the study team members to convey the content of the [Reference as Required: Statistical Analysis Complete (SAC)] deliverable.

**RAP Author(s):**

| Approver | Date |
|---|---|
| PPD<br>Principal Statistician (Future Pipelines Discovery, Clinical Statistics) | 7-May-2020 |

**RAP Team Reviews:**

| Reviewer | Date | RAP Team Review Confirmations |
|---|---|---|
| PPD<br>Director, Clinical Development | 8-May-2020 | E-mail |
| PPD<br>Medicine Development Leader, Development | 8-May-2020 | E-mail |
| PPD<br>Director Discovery Medicine Scientist, PCPS | 8-May-2020 | E-mail |
| PPD<br>Director, Patient Reported Outcomes<br>Patient Focused Outcomes | 7-May-2020 | E-mail |
| PPD<br>Principal Programmer, Future Pipelines Discover Clinical Programming | 8-May-2020 | E-mail |
| PPD<br>Director, Clinical Pharmacology Modeling and Simulation, CPMS | 8-May-2020 | E-mail |

**Clinical Statistics and Clinical Programming Line Approvals:**

| Approver |
|---|
| PPD<br>Director, Future Pipelines Discovery, Clinical Statistics |
| PPD<br>Programming Manager, Future Pipelines Discovery Clinical Programming |

## TABLE OF CONTENTS

PAGE

# 1. INTRODUCTION

The purpose of this reporting and analysis plan (RAP) is to describe the analyses to be included in the Clinical Study Report for Protocol:

| Revision Chronology: | | |
|---|---|---|
| 201000 | 6-Aug-2016 | Original |
| 201000/ 01 | 15-Nov-2016 | To clarify requirements for liver safety monitoring and stopping criteria, as well as other administrative changes. |
| 201000 / 02 | 07-Dec-2017 | To increase access to the trial for participants who will more closely reflect the intended treatment population of PBC patients while maintaining safety, to clarify some existing criteria and information, to make administrative changes, and to correct minor typographical and grammatical errors |

# 2. SUMMARY OF KEY PROTOCOL INFORMATION

## 2.1. Changes to the Protocol Defined Statistical Analysis Plan

There were no changes or deviations to the originally planned statistical analysis specified in the protocol amendment 2 [(Dated: 07/Dec/2017)].

## 2.2. Study Objective(s) and Endpoint(s)

| Objectives | Endpoints |
|---|---|
| **Primary** | |
| • To investigate the dose response of oral GSK2330672 on itch in PBC patients with moderate to severe pruritus at Baseline. | • Mean change from Baseline at Week 16 in the Mean Worst Daily Itch Score[1]. |
| **Key Secondary** | |
| • To characterize the effects of GSK2330672 compared to placebo on impact of symptoms and quality of life in PBC patients with moderate to severe pruritus at Baseline. | • Mean change from Baseline at Week 16 in PBC-40 Scale.* |
| • To evaluate the effects of | In participants meeting the criteria for |

| Objectives | Endpoints |
|---|---|
| GSK2330672 compared to placebo on markers of disease among participants at high risk of PBC progression (i.e., those with serum ALP concentrations ≥1.67xULN and/or total bilirubin concentrations >ULN at Day 1). | high risk of PBC progression:<br>• Mean change from Baseline at Week 16 in serum ALP concentrations.<br>• Proportion of participants having serum ALP concentrations <1.67x ULN and total bilirubin concentrations ≤ULN at Week 16.<br>• Mean change from Baseline at Week 16 in serum ALT, AST, GGT, total bilirubin and albumin concentrations and PT/INR. |
| **Other Secondary** | |
| • To determine the safety and tolerability of GSK2330672 compared to placebo in PBC patients with moderate to severe pruritus at Baseline. | • AEs, clinical laboratory parameters, ECGs, vital signs and the GSRS. |
| • To evaluate the effects of GSK2330672 compared to placebo on itch response rates in PBC patients with moderate to severe pruritus at Baseline. | • Proportion of participants who are responders at Week 16 based on each of the following separate definitions:<br>  • Mean Worst Daily Itch Score[1] of <4.<br>  • At least a 30% reduction from Baseline in the Mean Worst Daily Itch Score[1].<br>  • At least a 2-point reduction from Baseline in the Mean Worst Daily Itch Score[1].<br>• Mean number of responder days during Weeks 5 to 16 based on each of the following separate responder day[2] definitions:<br>  • Worst Daily Itch Score of <4.<br>  • Worst Daily Itch Score at least 30% lower than the Baseline Mean Worst Daily Itch Score.<br>  • Worst Daily Itch Score at least 2-points lower than the Baseline Mean Daily Score.+ |

| Objectives | Endpoints |
|---|---|
| • To further characterize the effects of GSK2330672 compared to placebo on symptoms and quality of life in PBC patients with moderate to severe pruritus at Baseline. | Change from Baseline at Week 16 in:<br>• Mean Daily Sleep Score[3].<br>• Mean Daily Fatigue Score[4].<br>• 5-D Itch Scale. |
| • To evaluate the effects of GSK2330672 compared to placebo on total serum bile acid concentrations and on bile acid synthesis in PBC patients with moderate to severe pruritus at Baseline. | • Mean change from Baseline at Week 16 in serum concentrations of total bile acids and serum C4 as a marker of bile acid synthesis. |
| • To investigate the PK of oral GSK2330672 in PBC patients with moderate to severe pruritus at Baseline. | • Plasma concentrations of GSK2330672 after sparse sampling (PK parameters will be reported if data permit). |
| **Exploratory** | |
| • To evaluate participant's treatment experience and health status with GSK2330672 compared to placebo in PBC patients with moderate to severe pruritus at Baseline. | • Treatment benefits and disadvantages as elicited in a Participant Treatment Experience Assessment at Week 16.<br>• Time to worsening of itch during Weeks 17-20 in participants with an improved Worst Daily Itch Score at Week 16 relative to Baseline.<br>• Change from Baseline at Week 16 in EQ-5D-5L health dimensions and utility index.<br>• Change from Baseline at Week 16 in Patient Global Impression of Severity (PGI-S) and Patient Global Impression of Change (PGI-C). |
| • To evaluate the effect of GSK2330672 compared to placebo on exploratory biomarkers of PBC and bile acid physiology in PBC patients with moderate to severe pruritus at Baseline. | • Change from Baseline at Week 16 in serum autotaxin, fibroblast growth factor-19 (FGF-19), enhanced liver fibrosis (ELF) test and individual serum bile acid species. |
| • To evaluate the effect of GSK2330672 compared to placebo on serum lipids and absorption of fat-soluble vitamins | • Change from Baseline at Week 16 in fasting lipid profile, including direct low density lipoprotein (LDL) |

| Objectives | Endpoints |
|---|---|
| in PBC patients with moderate to severe pruritus at Baseline. | cholesterol.<br>• Change from Baseline at Week 16 in Vitamins A, D, E and K. |
| • To evaluate the effect of GSK2330672 compared to placebo on depressive symptoms associated with PBC in PBC patients with moderate to severe pruritus at Baseline. | • Change from Baseline at Week 16 in the Beck Depression Inventory-II (BDI-II). |
| • To explore the effect of GSK2330672 on ursodeoxycholic acid (UDCA) concentrations in PBC patients with moderate to severe pruritus at Baseline. | • Plasma concentrations of UDCA after sparse sampling (in participants on UDCA). |
| • To evaluate the pharmacogenomics of PBC and response to GSK2330672 in PBC patients with moderate to severe pruritus at Baseline. | • Pharmacogenomics for genes related to PBC, pruritus and IBAT response (in consenting participants). |
| • To evaluate the potential of actigraphy to quantify scratching activity and sleep quality due to pruritus associated with PBC and to assess the effect of GSK2330672 on these parameters in PBC patients with moderate to severe pruritus at Baseline. | • Mean scratching event duration, and frequency (events/night) of specific scratching movements measured by wrist actigraphy and derived parameters (in participants in the Actigraphy Sub-study). |

1. Mean Worst Daily Itch Score: participant's itch severity is recorded on an electronic diary (eDiary) each morning and evening using a 0-10 numerical rating scale (NRS). The worst of these 2 scores is the Worst Daily Itch score. The Mean Worst Daily Itch Score is the worst daily itch score averaged over the 7 days preceding randomization at Visit 3 (Baseline) and over the 7 days preceding Visit 6 (Week 16).
2. A responder day is based on the Worst Daily Itch Score recorded using the 0-10 NRS for that day. When the responder day definition is relative to Baseline, the Worst Daily Itch Score is compared to the Mean Worst Itch Daily Score for Baseline (see 1 above).
3. Mean Daily Sleep Score: participant's sleep quality is recorded on the eDiary each morning using a 0-10 NRS and the Daily Sleep Score is averaged over the 7 days preceding randomization at Visit 3 (Baseline) and over the 7 days preceding Visit 6 (Week 16).
4. Mean Daily Fatigue Score: participants fatigue level is recorded on the eDiary each evening using a 0-10 NRS and the Daily Fatigue Score is averaged over the 7 days preceding randomization at Visit 3 (Baseline) and over the 7 days preceding Visit 6 (Week 16).

   Note:

   For all other endpoints Baseline is defined as the Visit 3 assessment and Week 16 is defined as the Visit 6 assessment.

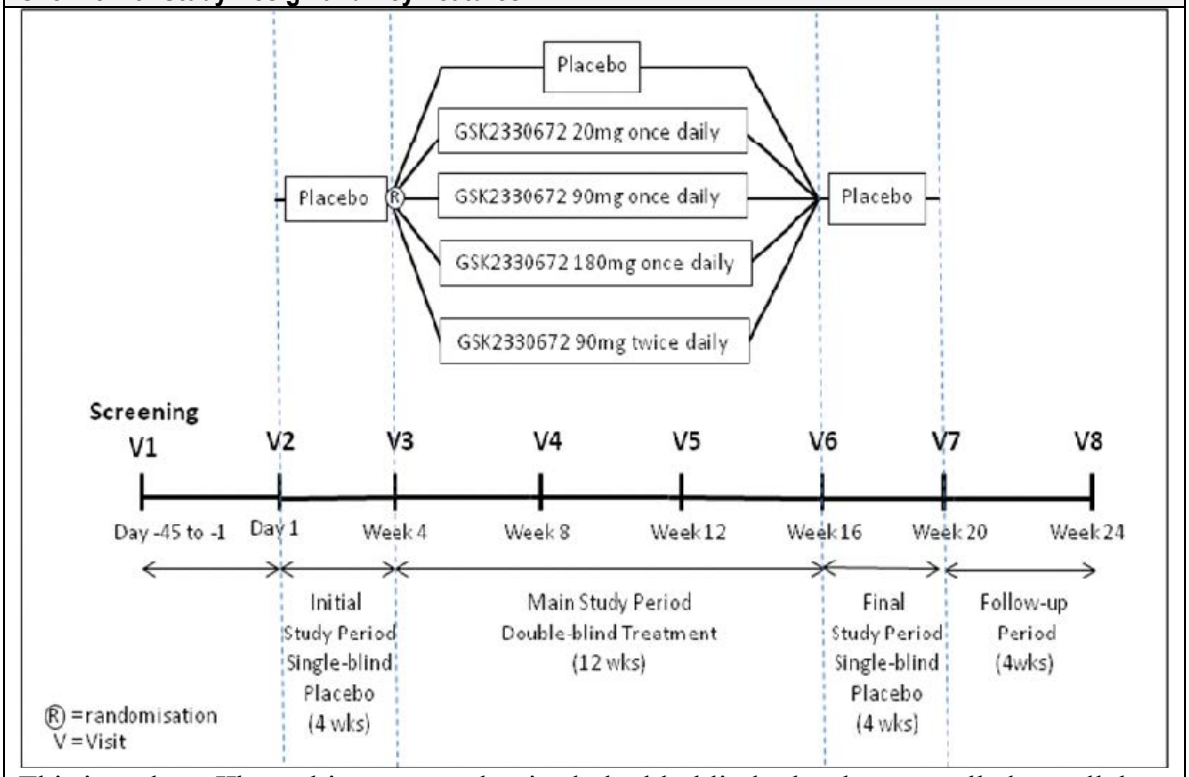
   In addition to Week 16, which is the primary time point of interest, other intermediate time points will also be assessed.

   *Mean change from Baseline at Week 16 in all the domains of PBC-40 Scale.

   +Mean Daily Score is a comparison to the 7 day mean worst daily itch score at baseline.

## 2.3. Study Design

**Overview of Study Design and Key Features**




This is a phase IIb, multicenter, randomised, double-blind, placebo-controlled, parallel group, dose-finding study in adults with moderate to severe pruritus associated with PBC. The study has an adaptive design which aims to utilize interim data to further inform and potentially optimize the doses under investigation (see Section 10.3.4 of the protocol).

Following the Screening Visit, there will be four study periods:

• Initial Study Period: Single-blind placebo treatment for 4 weeks during which participants' symptoms will be recorded in the eDiary to establish baseline symptoms and assess eligibility for randomization and compliance with study procedures.

• Main Study Period: Eligible participants will be randomised to receive 12 weeks double-blind treatment with either placebo or one of 4 dose regimens of GSK2330672 (20 mg, 90 mg or 180 mg taken once daily or 90 mg twice daily). Randomization will be stratified based on participant's risk of PBC disease progression based on serum ALP and total bilirubin concentrations at Day 1 (see Section 7.3.1).

• Final Study Period: Single-blind placebo treatment for 4 weeks to assess symptoms and safety post-completion of double-blind treatment.

• Follow-up Period: to assess symptoms and use of anti-pruritus medications by telephone visit.

| **Overview of Study Design and Key Features** | |
|---|---|
| Eligible participants may continue to receive some therapies for the treatment PBC provided these are maintained at stable doses and there is no plan to discontinue them during the study. Concomitant use of cholestyramine, colesevelam, colestipol or colestimide is not permitted until after completion of the Main Study Period. Obeticholic acid use is not permitted at any time during the study.<br><br>All study treatment will be administered orally twice a day, using placebo tablets as necessary to blind dose and regimen. Participants will commence single-blinded study treatment at Visit 2 [Day1]. During the Initial, Main and Final Study Periods it is important that participants remain blind to changes in study treatment from single-blind placebo to randomised double-blind treatment and subsequently to single-blind placebo. Investigator and study site staff communication with participants should ensure maintenance of each participant's blinding to treatment throughout the study.<br><br>A participant's total duration in the study will be approximately 24 weeks from Day 1 to completion of the Follow-up Telephone Visit.<br><br>The study also includes an exploratory Actigraphy Sub-study which will be conducted at a limited number of study sites and in a sub-set of participants who give their consent. | |
| **Design Features** | This is a Phase IIb, multicentre, randomised, double-blind, placebo-controlled, parallel group, dose-finding study in adults with moderate to severe pruritus associated with PBC.<br><br>There are 3 core study periods; Each participant will receive<br><br>1) Single-blind placebo for 4 weeks (Initial study period - ISP),<br>a. Starting at Visit 2 (Day 1) and ending on Visit 3<br>2) Double-blind study treatment for 12 weeks (Main study period - MSP),<br>a. Starting at Visit 3 (Baseline) and ending at Visit 6<br>3) Single-blind placebo for a further 4 weeks (Final study period - FSP).<br>a. Starting at Visit 6 and ending at Visit 7<br>With a follow-up phone call at Visit 8<br><br>This study has an adaptive design, which aims to utilise interim data to further inform and potentially optimise the doses under investigation. |
| **Dosing** | The study compound is GSK2330672.<br>Each subject will be administered study treatment orally twice a day, using |

| Overview of Study Design and Key Features | |
|---|---|
| | placebo tablets as necessary to blind dose and regimen.<br>During the main study period, participants will be randomised to receive 12 weeks double-blind treatment of one of the following dose regimens:<br>• placebo<br>• 20 mg once daily (QD)<br>• 90 mg once daily (QD)<br>• 180 mg once daily (QD)<br>• 90 mg twice daily (BID)<br>• In all QD dose groups in order to maintain the blind randomised treatment will comprise GSK2330672 taken for the morning dose and placebo taken for the evening dose. |
| **Time & Events** | • [Refer to Appendix 2: Schedule of Activities] |
| **Treatment Assignment** | Participants for whom exclusion/inclusion criteria are satisfied, and the randomisation criteria at the end of the ISP are met, will be randomised to double-blind treatment.<br>• Randomisation will be stratified by region (Europe and Australia, North America, Japan) and participant's risk of PBC disease progression (Low, High – see Table 1). They will be randomly allocated to a dose group as per the ratios within each disease progression risk stratum. The same randomisation ratios will be used within each region. The different randomisation ratio in the high risk group is to allow for increased precision of comparisons between the high dose group and placebo in analysis of the high risk population. It is believed that a subject's itch will not differ between the Risk Group and Region strata. There will be a check of whether risk group affect the results and if so the bias associated with ignoring this effect will be estimated. |
| **Interim Analysis** | • Interim analyses took place after approximately 40 participants reached Visit 5 (Week 12). |

## 2.4. Statistical Hypotheses / Statistical Analyses

The primary objective is to estimate the relationship between doses of GSK2330672 and Itch response. The primary null hypothesis ($H_0$) assumes that there is no effect of the test drug at any dose level and therefore no evidence of dose response. The alternate hypothesis ($H_1$) assumes that there is a dose response with GSK2330672 on participant experience of itch, where there is a better improvement in itch with increasing dose.

# 3. PLANNED ANALYSES

## 3.1. Interim Analyses

An interim analysis (IA) of primary efficacy and safety was performed as of the June of 2018. Based on the IA results the dose levels of GSK2330672 included at the randomisation was modified, in addition to the total sample size of the study was increased to a total of 140 randomised participants.

Please refer to the IA RAP for further details.

## 3.2. Final Analyses

The final planned primary analyses will be performed after the completion of the following sequential steps:

1. All participants have completed the study as defined in the protocol. The study is completed after all participants have completed the follow-up telephone interview (Visit 8, Week 24).
2. All required database cleaning activities have been completed and final database release and Data Base Freeze (DBF) has been declared by Data Management.
3. All criteria for unblinding the randomization codes have been met.
4. Randomization codes have been distributed according to RandAll NG procedures.

# 4. ANALYSIS POPULATIONS

| Population | Definition / Criteria | Analyses Evaluated |
|---|---|---|
| Enrolled | All participants who sign the ICF | Screening Failure<br>Randomisation Failure<br>Safety |
| Randomised | All participants who were randomised to receive study treatment, regardless of whether they took randomised study treatment. | Study Population<br>Participant disposition summaries |
| ITT | The Intent-to-Treat (ITT) Population will comprise all randomised participants who receive at least one dose of randomised study treatment, have a baseline and at least one on-treatment assessment in the Main Study Period. Participants in the ITT Population will be classified according to the treatment as randomised. | Efficacy |
| PP | The Per Protocol (PP) Population is a subset of the ITT Population who adhered to the major | Efficacy (Sensitivity of |

| Population | Definition / Criteria | Analyses Evaluated |
|---|---|---|
| | protocol requirements. The protocol deviation criteria are provided in Section 4.1.. Participants to be excluded from the PP Population will be identified prior to unblinding of the data at the end of the study. The PP Population may not be analysed if each Treatment Group within the PP population comprises ≥ 80% of the Treatment Group within the ITT Population | Primary endpoint analyses) |
| Safety | All randomised participants who take at least 1 dose of randomised study treatment. Participants will be analysed according to the treatment they actually received. | Safety |
| Initial | All participants who were enrolled, passed screening and entered the placebo run-in phase, taking at least one dose of placebo during this period. | |
| High Risk | This is a subset of the ITT population who are assigned to the High Risk strata for randomisation (based upon serum ALP concentrations ≥1.67xULN and/or total bilirubin concentrations >ULN at Day 1 (Visit 2). | Secondary Efficacy ALP, AST, ALT, GGT, total bilirubin, albumin, PT/INR |
| Restricted High Risk | This is a subset of the High Risk population. This population consists of all those assigned to the High Risk strata for randomisation who meet the ALP/bilirubin criteria at both Visit 2 and Visit 3 (serum ALP concentrations ≥1.67xULN and/or total bilirubin concentrations >ULN at V2 and V3). | Secondary Efficacy Supplementary Analysis |
| Responders (>=2 point improvement) | This is a subset of the ITT population. Those whose mean worst daily itch score improves by 2 or more at Week 16 (as compared to Baseline). | Time to Worsening Analysis |
| Actigraphy | The subset of the ITT population who gave additional consent for and performed measurements of Actigraphy | Actigraphy |
| PK | Any randomised participant who had at least 1 PK sample | PK population |

Refer to Appendix 12: List of Data Displays which details the population used for each display.

## 4.1. Protocol Deviations

Important protocol deviations (including deviations related to study inclusion/exclusion criteria, conduct of the trial, patient management or patient assessment) will be summarised and listed.

Important deviations which result in exclusion from the analysis population will also be summarised and listed. (Please refer to Appendix 1: Protocol Deviation Management and Definitions for Per Protocol Population).

Protocol deviations will be tracked by the study team throughout the conduct of the study in accordance with the Protocol Deviation Management Plan [April 16, 2020 Version 3.0].

If a Per Protocol Population is being defined: Important deviations which result in exclusion from the analysis population will also be summarised and listed. (Please refer to Appendix 1: Protocol Deviation Management and Definitions for Per Protocol Population).

- Data will be reviewed prior to unblinding and freezing the database to ensure all important deviations and deviations which may lead to exclusion from the analysis are captured and categorised on the protocol deviations dataset
- Protocol deviations will be tracked by the study team throughout the conduct of the study in accordance with the study specific Protocol Deviation Management Plan, which is specifically developed for this study.
- Important protocol deviations (which include deviations related to study inclusion/exclusion criteria, conduct of the trial, patient management or patient assessment) will be summarised and listed. This will be based on data recorded in the eCRF, particularly the protocol deviations page.
- Important deviations which result in exclusion from the PP will also be summarised and listed.
- The study endpoints will be reported using the populations detailed in this Section of this document.
- This dataset will be the basis for the summaries and listings of protocol deviations.

A separate summary and listing of all inclusion/exclusion criteria deviations will also be provided. This summary will be based on data as recorded on the inclusion/exclusion page of the eCRF.

# 5. CONSIDERATIONS FOR DATA ANALYSES AND DATA HANDLING CONVENTIONS

## 5.1. Study Treatment & Sub-group Display Descriptors

| Treatment Group Descriptions | | | |
|---|---|---|---|
| [RandAll NG / FSO Randomization System] | | Data Displays for Reporting | |
| Code | Description | Description | Order in TLF |
| AA | Placebo for 16 weeks | Placebo | 1 |
| BB | GSK2330672 20 mg once daily for 12 weeks followed by placebo for 4 weeks | GSK2330672 20 mg QD | 2 |
| CC | GSK2330672 90 mg once daily for 12 weeks followed by placebo for 4 weeks | GSK2330672 90 mg QD | 3 |
| DD | GSK2330672 180 mg once daily for 12 weeks followed by placebo for 4 weeks | GSK2330672 180 mg QD | 4 |
| EE | GSK2330672 90 mg twice daily for 12 weeks followed by placebo for 4 weeks | GSK2330672 90 mg BID | 5 |

Note: Refer to Appendix 14.

## 5.2. Baseline Definitions

For all endpoints (except those collected on the eDiary) the baseline will be the assessment performed at Visit 3 which is conducted prior to the first dosing of randomised medication that evening. Where there is more than one value at any given time point, the mean of the replicate assessments will be used as the value for that time point.

The assessment performed at Visit 3 for laboratory data, ECGs, and vital signs will be used as the baseline value in all data displays.

Where the baseline (Visit 3) data are missing for laboratory results, ECGs, and vital signs, they will be replaced by the last measurement prior to the first randomised dose (Visit 2 or Visit 1, excluding unscheduled visits).

For further information on deriving the endpoints, including rules for missing data, see Appendix 6 and Appendix 7.

Weekly assessments collected on the eDiary: The baseline GSRS value will be the most recent GSRS assessment completed by the participant prior to randomisation, similarly for the baseline GI symptom questions.

Daily data collected on the eDiary: the baseline will be as follows.

- **Mean Worst Daily Itch** will be the average of the maximum of the AM and PM scores in the 7 days prior to the baseline visit (Visit 3).

- **Mean Overall Daily Itch** will be the average of the mean of the AM and PM scores in the 7 days prior to the baseline visit (Visit 3).

- The **mean sleep score**, and **mean fatigue score** at baseline will be the average of the sleep scores and fatigue scores, respectively, recorded each day over the 7 days prior to the baseline visit (Visit 3).

Other itch symptom based eDiary assessments using a categorical response will employ a baseline value based on the average of the scores in the 7 days prior to the baseline visit, where the questions will be scored according to the process set out in Section 12.6.2. The categorical questions will also be summarised in terms of the mode of the responses in the 7 days prior to the visit, and the baseline for these summaries will be the mode of the responses in the 7 days prior to the baseline visit.

| Parameter | Study Assessments Considered as Baseline | | | | Baseline Used in Data Display |
|---|---|---|---|---|---|
| | Visit 3 (Pre-Main study dose) | Mean of 7 days prior to Visit 3 | Completed during Week preceding the Baseline Visit | Mode of response in 7 days prior to Visit 3 | |
| **[Efficacy]** | | | | | |
| Mean Worst Daily Itch,<br>Mean sleep score, Mean fatigue score<br>Mean overall itch score | | X | | | Baseline |
| Other eDiary questions | | X | | X | Baseline |
| PBC-40<br>5-D Itch,<br>PGI-S,<br>EQ-5D-5L,<br>BDI-II | X | | | | Visit 3 (V3) |
| GSRS | | | X | | Baseline |

| Parameter | Study Assessments Considered as Baseline | | | | Baseline Used in Data Display |
|---|---|---|---|---|---|
| | Visit 3 (Pre-Main study dose) | Mean of 7 days prior to Visit 3 | Completed during Week preceding the Baseline Visit | Mode of response in 7 days prior to Visit 3 | |
| GI Symptoms Questions | | | X | | Baseline |
| **[Safety]** | | | | | |
| laboratory data, ECGs, and vital signs | X | | | | Visit 3 (V3) |
| PD/Biomarker samples:<br>Serum autotaxin,<br>FGF-19<br>ELF<br>Serum bile acid species<br>C4 | X | | | | Visit 3 (V3) |

## 5.3. Multicentre Studies

- In this multicentre global study, enrolment will be presented by Region.
- Key summary tables and figures will be presented by Region, where appropriate.
- Where numbers allow further key summary tables may be presented by country.

| Region | Countries |
|---|---|
| North America | USA, Canada |
| Japan | Japan |
| Europe and Australia | Spain, UK, France, Australia, Italy, Germany, and Poland |

**NOTES: Japan**

Key summary tables and figures will be presented for the participants from Japan only vs the rest of the world to determine whether the response in these participants is similar to those in the study as a whole. The efficacy figures for Region should display the overall results from the modelling, along with the data for Japan.

## 5.4. Examination of Covariates, Other Strata and Subgroups

### 5.4.1. Covariates and Other Strata

The list of covariates and other strata may be used in descriptive summaries and statistical analyses, some of which may also be used for subgroup analyses. Additional covariates and other strata of clinical interest may also be considered.

No formal statistical analysis will be produced if number of participants per cell is small for any of the categories listed below and only summary tables and listing will be produced.

| Category | Details |
|---|---|
| Covariates to be included in all models which consider change from baseline | |
| Centred Mean Worst Daily Itch at Baseline | Continuous |
| Covariates that will be investigated as part of supplementary analysis of the primary endpoint | |
| ALP (in units of ULN) at Day 1 | Continuous |
| ALP (in units of ULN) at Baseline | Continuous |
| Total Bilirubin (in units of ULN) at Day 1 | Continuous |
| Total Bilirubin (in units of ULN) at Baseline | Continuous |
| Sex | Male<br>Female |
| UDCA | Patient on UDCA at Baseline<br>Patient not on UDCA at Baseline |
| serum bile acids at Baseline | Continuous |
| C4 at Baseline | Continuous |
| Time from diagnosis | Continuous<br>Note: Based on date of diagnosis and baseline date |
| Beck Depression Inventory at Baseline | Continuous |
| Age at diagnosis | Continuous<br>Note: based on date of diagnosis and DOB (date of onset of symptoms/diagnosis up to screening) |
| Baseline Itching Group | <4 'Mild' |

| Category | Details |
|---|---|
| | ≥4 - <7 'Moderate'<br>≥7 - =10 'Severe' |
| Previous therapy | None<br>Cholestyramine<br>Cholestyramine refractory |
| Concomitant background therapy | The following is a list of the concomitant background therapy:<br>- Rifampicin<br>- Oral opiate antagonists (naltrexone or nalmefene)<br>- Sertraline<br>- Gabapentin<br>- Antihistamines<br>- Nalfurafine |
| Concomitant background therapy (Yes vs. No) | Yes<br>No<br><br>Note: if a participant used any of the above listed concomitant background therapy then yes, otherwise it is no. |

### 5.4.2. Examination of Subgroups

The list of subgroups may be used in descriptive summaries and statistical analyses. Additional subgroups of clinical interest may also be considered.

- If the percentage of participants is small within a particular subgroup, then the subgroup categories may be refined prior to unblinding the trial.
- If the category cannot be refined further, then descriptive rather than statistical comparisons may be performed for the particular subgroup.

| Subgroup | Categories |
|---|---|
| Country | Japan only<br>The rest of the world |
| Strata to be investigated for impact on primary analysis | |

| Subgroup | Categories |
|---|---|
| Risk of Disease Progression | Low<br>High |
| Region | USA<br>Japan<br>Europe and Australia |

## 5.5. Multiple Comparisons and Multiplicity

Analyses of other efficacy endpoints will not be participant to any multiplicity adjustment.

## 5.6. Other Considerations for Data Analyses and Data Handling Conventions

Other considerations for data analyses and data handling conventions are outlined in the appendices:

| Section | Component |
|---|---|
| 12.1 | Appendix 1: Protocol Deviation Management and Definitions for Per Protocol Population |
| 12.2 | Appendix 2: Time & Events (Schedule of Activities) |
| 12.3 | Appendix 3: Assessment Windows |
| 12.4 | Appendix 4: Study Phases and Treatment Emergent Adverse Events |
| 12.5 | Appendix 5: Data Display Standards & Handling Conventions |
| 12.6 | Appendix 6: Derived and Transformed Data |
| 12.7 | Appendix 7: Reporting Standards for Missing Data |
| 12.8 | Appendix 8: Values of Potential Clinical Importance |
| 12.9 | Appendix 9: Population Pharmacokinetic (PopPK) Analyses |
| 12.10 | Appendix 10: Pharmacokinetic / Pharmacodynamic Analyses |
| 12.11 | Appendix 11: Abbreviations & Trade Marks |
| 12.12 | Appendix 12: List of Data Displays |
| 12.13 | Appendix 13: Example Mock Shells for Data Displays |
| 12.14 | Appendix 14: Interim Analysis Results and update to RAP |

# 6. STUDY POPULATION ANALYSES

## 6.1. Overview of Planned Study Population Analyses

The study population analyses will be based on the most appropriate population, unless otherwise specified.

Study population analyses including analyses of participant's disposition, protocol deviations, demographic and baseline characteristics, prior and concomitant medications, and exposure and treatment compliance will be based on GSK Core Data Standards. Details of the planned displays are presented in Appendix 12: List of Data Displays.

# 7. EFFICACY ANALYSES

## 7.1. Primary Efficacy Analyses

### 7.1.1. Endpoint / Variables

Mean change from baseline Week 4 (V3) at Week 16 (V6) in Mean Worst Daily Itch Score defined as:

- Difference from baseline.

The difference from baseline will be the primary endpoint at Visit 6 for the Mean Worst Daily Itch. The final chosen model will be presented in terms of the LS Mean change from baseline for each treatment group/total daily dose in the study along with associated 95% CIs. The difference from placebo will also be presented with 95% CIs but as the study is exploratory, a p-value will not be presented.

The Centred Mean Worst Daily Itch at Baseline will be included in the models as a covariate. It will be calculated using only the participants with change from baseline in Mean Worst Daily Itch data available for analysis within the population of interest.

Comparison between each total daily dose and placebo will be provided using 2-sided t-test with alpha level of 0.05 on the mean change from baseline at Week 16 (V6) in Mean Worst Daily Itch score.

A plot of the fitted dose-response model and 95% prediction interval against total daily dose, with change from baseline at Week 16 overlaid as a scatter plot will be generated.

In addition, to the above dose-response model and above plot four dose-response models will be generated to produced four plots for model checking and supplementary analyses as following:

1. A plot of the fitted dose-response model and 95% prediction interval against total daily dose, with change from baseline at Week 16 overlaid as a scatter plot using participants on the QD dose levels only

2. A plot of the fitted dose-response model and 95% prediction interval against total daily dose, with change from baseline at Week 16 overlaid as a scatter plot using participants who are on the BID dose levels only

3. A plot of the fitted dose-response model and 95% prediction interval against total daily dose, with change from baseline at Week 16 overlaid as a scatter plot overlay of the QD dose levels and BID dose levels on one plot (overlay plot 1 and 2 above)

4. A plot of the fitted dose-response model and 95% prediction interval against total daily dose, with change from baseline at Week 16 overlaid as a scatter plot is converting the 180 QD dose level to the 90 BID dose level and using the combined dose levels on one plot.

See Appendix 14 for further details.

### 7.1.2. Summary Measure

Summary measures are Daily Itch Scores (AM, PM and Worst Daily), Mean Worst Daily Itch, and Change from Baseline in Mean Worst Daily Itch at Week 16 (V6).

### 7.1.3. Population of Interest

The primary efficacy analyses will be based on the ITT population. A sensitivity analysis of the results will be performed using the PP population.

### 7.1.4. Statistical Analyses / Methods

Details of the planned displays are provided in Appendix 12: List of Data Displays and will be based on GSK data standards and statistical principles.

Unless otherwise specified, endpoints / variables defined in Section 7.1.1 will be summarised using descriptive statistics, graphically presented (where appropriate) and listed.

#### 7.1.4.1. Statistical Methodology Specification

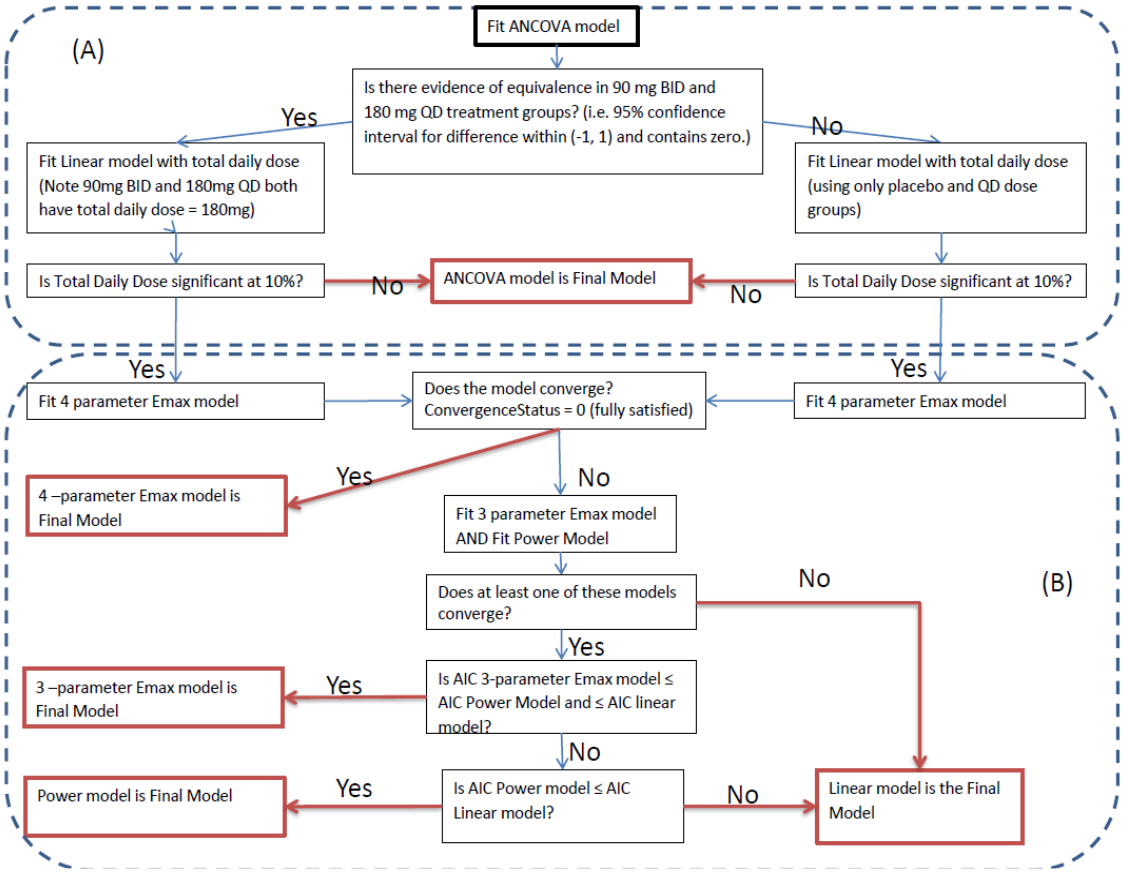
| **Endpoint / Variables** |
|---|
| Change from baseline (V3) at Week 16 (V6) in Mean Worst Daily Itch defined as:<br>• Difference from baseline. |
| **Model Specification** |

The dose-response of the change from baseline in mean worst daily itch will be analysed as outlined in Figure 1. The modelling process is split into two sections:

(A) Preliminary checks on 1) whether there is a dose-response relationship and 2) whether to include BID group in dose-response modelling

(B) Dose-response modelling, identifying the most appropriate model

**Figure 1: Flow chart of the model fitting process (see Appendix 14 for more details related to Figure 1)**




This process, starting from the linear model in (A) and comparing models in (B) (i.e. assuming the same decision as to whether to exclude or include the BID group).

## 1. Analysis of Covariance (ANCOVA)

ANCOVA: Preliminary check of whether to include BID group in dose-response modelling

- An ANCOVA model will be fitted using all treatment groups to test whether 90 mg BID and 180 mg QD treatment groups are equivalent with no significant difference.
- The model should include:

  Trtcd - Treatment Group – coded as:
  AA - placebo, BB – 20 mg QD, CC – 90 mg QD, DD – 180 mg QD, EE – 90 mg BID.

Cbl_itch - Centred Mean Worst Daily Itch at Baseline

- If the confidence interval for the difference between 90 mg BID and 180 mg QD lies within (-1, 1) and includes 0 then all dose groups will be included in the following models (where 90 mg BID and 180 mg QD are both total daily dose of 180 mg). Otherwise, the 90 mg BID group will be excluded from the dose-response analysis.
- If ANCOVA is chosen as the final model (i.e. there is no evidence of a dose-response relationship) the MED will be the lowest dose at which the change from baseline in Mean Worst Daily Itch is < -2. If there are no treatment groups that achieve a change from baseline of < -2 this will be presented as NA in the appropriate table.

**2. Linear Model**

Linear Model: Check whether there is a dose-response relationship

- The linear model will be fitted to the data to test whether there is evidence of a dose-response relationship.
  - $Y = \alpha + \beta * Cbl\_itch + \gamma * Dose$
- Whether all treatment groups are used or just placebo and 3 QD treatment groups will depend on the results of the ANCOVA model
- The model should include:

  Dose - Total daily dose – continuous variable (e.g. 0, 20, 90 and 180 mg).

  Cbl_itch - Centred Mean Worst Daily Itch at Baseline

- If total daily dose is significant in the model at 10% then proceed with non-linear dose-response modelling, otherwise the ANCOVA model should be reported as the best model.
- The MED will be calculated for the linear model as:

$$MED = \frac{-2 - \alpha}{\gamma}$$

**3. Dose-response Modelling**

- The dose-response modelling process will attempt to fit non-linear models to the data and using decision criteria (as outlined in Figure 1 above) a final model will be chosen.

  - Fit 4-parameter Emax model
    - If converges – 4-parameter Emax model is final chosen model
    - If doesn't converge then fit 3-parameter Emax model and Power model
      - If there are convergence problems with both of these models
        - ➢ Take Linear Model as final chosen model

- Otherwise

  ➢ Take the final chosen model to be the one with the lowest AIC (Akaike's Information Criterion) (linear, power or 3-parameter Emax models)

- All of the non-linear models will include:

  Dose - Total daily dose – continuous variable (e.g. 0, 20, 90 and 180 mg).

  Cbl_itch - Centred Mean Worst Daily Itch at Baseline

- Whether all treatment groups are used or just placebo and 3 QD treatment groups will depend on the results of the ANCOVA model.

**4. <u>4-Parameter Emax Model</u>**

Dose-response Modelling: 4-Parameter Emax Model

The 4-parameter Emax model is given as:

$$Y = E_0 + Beta * Cbl\_Itch + \frac{Dose^{slope} * E_{max}}{Dose^{slope} + ED_{50}^{slope}}$$

Where $E_0$ is the placebo effect and $E_{max}$ is the asymptotic maximum change from placebo effect. $ED_{50}$ is the dose at which the response is mid-way point between $E_0$ and $E_{max}$.

Centred Mean- Mean Worst Daily Itch at Baseline

(Cbl_itch) will be fitted as a covariate.

The MED will be calculated as:

$$MED = \frac{ED_{50}}{\left(\frac{E_{max}}{[-2 - E_0]} - 1\right)^{1/slope}}$$

**5. <u>3-Parameter Emax Model</u>**

Dose-response Modelling: 3-Parameter Emax Model

The 3-parameter Emax model is given as:

$$Y = E_0 + Beta * Cbl\_Itch + \frac{Dose * E_{max}}{(Dose + ED_{50})}$$

This is equivalent to the 4-parameter Emax model with slope fixed and equal to 1. Mean- Centred Mean Worst Daily Itch at Baseline will be fit as a covariate in the model.

The MED is calculated as:

$$MED = \frac{ED_{50}}{\left(\frac{E_{max}}{[-2 - E_0]} - 1\right)}$$

6. **Power Model**

Dose-response Modelling: Power Model

As per Kirby et al (2011) the Power Model will take the form:

$$Y = E_0 + Beta * Cbl\_Itch + slope * dose^{lambda}$$

Note that unlike the power model commonly used in PK dose proportionality, this model includes an intercept and will also include the Centred Mean Worst Daily Itch at Baseline as a covariate.

The MED will be calculated as:

$$MED = \left(\frac{-2 - E_0}{slope}\right)^{1/lambda}$$

**Model Checking & Diagnostics**

**For all endpoints that uses Non-linear modelling and ANCOVA:**

Distributional assumptions underlying the model used for analysis 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 departures from the distributional assumptions, alternative models will be explored using appropriate transformed data.

**For the Change from Baseline in Mean Worst Daily Itch endpoint using ANCOVA:**

To check whether the assumption that there is no impact of the deliberately imbalanced strata (i.e. that risk group does not affect the change from baseline in mean worst daily itch) a supplementary analysis will be undertaken to test whether there are differences between the strata (supplementary analysis: investigation of risk group within this section listed below)

**Model Results Presentation**

- For the final chosen model:
  - The LS mean change from baseline (and 95% CIs) as well as the mean differences in change from baseline (and 95% CIs) of each study total daily dose levels from placebo will be displayed along with the 95% CIs

in a table.
  - A figure will be presented for the LS mean difference from placebo.

- Where the final chosen model is an ANCOVA, these tables and figures will be presented for treatment group, including the 90 mg BID group. Otherwise they will be presented for the total daily dose at each dose level in the study (e.g. 20, 90 and 180 mg).

- Similar tables will be presented for the results of the supplementary analyses.

- The SAS model output, including estimated model parameters, will be provided in the listings

**Subgroup Analyses**

- The stratification variable Risk Group (ALP) will be fitted in the ANCOVA model as a covariate, as well as the interaction between Risk Group (ALP) and Treatment Group.

**Estimating Bias**

- If Risk Group (ALP) is found to be significant in the ANCOVA model an estimation of the bias in the results arising from ignoring risk group within the analysis will be presented.
- A table will present the estimated change from baseline for each treatment group for the ANCOVA excluding risk group and the ANCOVA including risk (ALP) group. The table will also show the difference from placebo for each treatment group, and difference between 90 mg BID and 180 mg QD.
- The bias will be estimated as:
  - Estimate excluding risk group (ALP) – estimate including risk group (ALP).

- A figure will also present the results.

- An overall estimate of the maximum bias associated with ignoring risk group will be made by averaging the bias of the 20 mg QD, 90 mg QD and 90 mg BID groups vs. placebo.

**Supplementary Analyses**

- MMRM will be fitted using all available mean worst daily itch scores at Visit 4, Visit 5, Visit 6 and Visit 7 to investigate the mean change from baseline in mean worst daily itch at each visit and mean difference from placebo with 95% CIs. This will include the baseline mean worst daily itch and a treatment by visit interaction. An unstructured covariance structure will be assumed; if this does not converge then an autoregressive [AR(1)] or compound symmetry (CS) will be selected based on the lowest Akaike's criterion.

- The final dose-response model for the change from baseline (i.e. linear, Emax, power from as described above) will be fitted using only those participants in the PP Population. The centered mean worst daily itch, as defined for the PP population will be included as a covariate.

- An ANCOVA model for the change from baseline will be fitted including each of the covariates listed in Section 5.4.1, one at a time, in addition to the treatment group and interaction with treatment group and results will be presented in tables.

## 7.2. Secondary Efficacy Analyses

### 7.2.1. Endpoint / Variables

- The secondary efficacy analysis outlined in this section are listed under Section 2.2. and covers end-points associated with responder definitions, patient reported outcomes (PROs), and the analyses in the High Risk population.

- The analysis for secondary endpoints that are defined as Pharmacokinetic (PK) or Pharmacodynamic (PD) biomarkers will be described in Section 9.1 and Section 10.1 respectively.

### 7.2.2. Summary Measure

**Responders in itch:**

Responders in itch are listed below and will be summarised accordingly:

1. **Proportion of participants who are responders in Mean Worst Daily Itch score at Week 16 (V6) on each of the following separate\ definitions:**
   Improvement[§]≥ 2
   Score < 4
   Percentage Improvement[§] ≥30%

   [§] Mean Worst Daily Itch Score: participant's worst itch severity is recorded on an electronic diary (eDiary) each morning and evening using a 0-10 numerical rating scale (NRS), the Worst Daily Itch Score is the average of the maximum of the AM and PM scores in the 7 days preceding randomization at Visit 3 (Baseline) and over the 7 days preceding Visit 6 (Week 16).

2. **Mean number of responder days during Weeks 5 to 16 based on each of the following separate responder day definition:**
   Change from baseline at Week 16 >= 2
   Score at week 16 < 4
   Percentage change from baseline at Week 16 >30%

Days Response in Worst Daily Itch score will be reported as mean number of days, SD, and percentage for each dose level in which responder criteria has been met during the main study period.

**PRO: Other eDiary Itch questions:**

In addition to the primary endpoint further questions in the eDiary capture information on the itch experienced by participants. Overall itch, sleep and fatigue use the same NRS (0-10), while the others are categorical.

**Summary measurement for Sleep, Fatigue, Overall Itch and other daily eDiary questions are:**

Mean Daily Fatigue Score

Mean Daily Sleep Score

**PRO: Validated Questionnaires:**

In addition to the eDiary there are a number of other PROs using validated questionnaires which aim to capture information on the subjects' impression of PBC (PBC-40), their itch (5-D itch), their quality of life (EQ-5D-5L), their mental wellbeing (BDI-II), their impression of severity of disease (PGI-S) and of change (PGI-C).

The Patient Treatment Experience Assessment (PTEA) collects information at Visit 6, or at early end of treatment visit, on how the participant felt about the study treatment.

The final follow-up phone-call (Visit 8) provides information on the participant's experience of itch at the end of the follow-up period.

**Other PRO measurements:**

5-D Itch (For each of the 5 domain scores of the 5-D Itch and the total 5-D Itch score)

PBC-40 domains (For each of the 6 domain scores of the PBC-40)

**High Risk of Disease Progression (ALP) :**

Participants are assigned to the High Risk strata at Visit 3 (baseline), for randomisation, based on ALP and total bilirubin measured in samples drawn at Visit 2 (ALP at V2 ≥ 1.67*ULN and/or total bilirubin at V2 >ULN).
As a result, the High Risk population will consist of all those assigned to the High Risk strata for randomisation.
The following secondary endpoints will be examined within the High Risk population:

- Mean change from Baseline at Week 16 in serum ALP concentrations.
- Proportion of participants having serum ALP concentrations <1.67x ULN and total bilirubin concentrations ≤ULN at Week 16.

- And in the Markers of Disease Measurements (Mean change from Baseline at Week 16 in serum ALT, AST, GGT, total bilirubin and albumin concentrations and PT/INR).

A sensitivity analysis of the results for the High Risk population will be carried out where the population is defined as all those assigned to the High Risk strata for randomisation, whose lab results at Visit 2 and Visit 3 confirm them as being at high risk of disease progression. This sensitivity analysis will only be carried out if the two populations differ.

**Disease Progression Risk Response Measurements:**

Disease Progression Risk Response

### 7.2.3. Population of Interest

The secondary efficacy analyses will be based on the ITT population, unless otherwise specified.

### 7.2.4. Statistical Analyses / Methods

Details of the planned displays are provided in Appendix 12: List of Data Displays and will be based on GSK data standards and statistical principles.

Unless otherwise specified, endpoints / variables defined in Section 7.2.1 will be summarised using descriptive statistics, graphically presented (where appropriate) and listed.

#### 7.2.4.1. Statistical Methodology Specification

**Endpoint / Variables**

Please see Section 2.2. for the list of the secondary endpoints.

**Model Specification**

The following is the list of the secondary endpoint and model specification:

| Endpoint | Endpoint Type | Population | Models | Covariates |
|---|---|---|---|---|
| **Responder** | | | | |
| Score < 4, Improvement ≥ 2 Improvement ≥ 30% | A | ITT | Logistic Regression | None |
| **Percentage of Responder Days** | | | | |
| Score < 4, Improvement ≥ 2 Improvement ≥ 30% | A | ITT | ANCOVA | None |
| | | | | |
| | | | | |
| **eDiary question and other PROs** | | | | |

| | | | | | |
|---|---|---|---|---|---|
| Mean Overall Itch | CFB | ITT | ANCOVA | Mean Overall Itch at Baseline[1] | |
| Mean Sleep | CFB | ITT | ANCOVA | Mean Sleep at Baseline[1] | |
| Mean Fatigue | CFB | ITT | ANCOVA | Mean Fatigue at Baseline[1] | |
| PBC-40 domain scores[2] | CFB | ITT | ANCOVA | PBC-40 domain scores at Baseline[1] | |
| 5-D Itch scores[3] | CFB | ITT | ANCOVA | 5-D Itch scores at Baseline[1] | |
| **Markers of Disease** | | | | | |
| Disease Progression Risk | A | High Risk | Logistic Regression | None | |
| ALP,<br>ALT<br>AST<br>GGT<br>Total Bilirubin<br>Albumin<br>PT/INR | CFB, | High Risk | ANCOVA | Marker at baseline[1] | |

A = Absolute
CFB = Change from baseline

[1] Baseline can be centred or not, it will not impact the results in the analysis
[2] For each PBC-40 domain
[3] For each 5-D Itch domain and the total 5-D itch score

*Within the Responders (>=2 point improvement) population only treatment groups with at least 8 participants will be included in the analysis.

**Model Checking & Diagnostics**

See under model checking & diagnostics outlines checks that should be carried out of the statistical assumptions made.

**Model Results Presentation**

- An ANCOVA model will be fitted to the data to investigate the difference in response over the dose-groups. The response will be the endpoint of interest (i.e. percentage of days, change from baseline in ALT), with the treatment group as the explanatory variable. Where appropriate the baseline will be included as a covariate (in analyses of change from baseline).

  A results table will present the LS mean of the endpoint of interest for each of the treatment groups (and 95% CIs) as well as the LS mean differences between each treatment and placebo and the mean difference between 180 mg QD and 90 mg BID treatment groups (and 5% CIs).

  The mean differences in change from baseline from placebo will be plotted along with the

| |
|---|
| corresponding 95% CIs in a Figure, with a reference line at y=0.<br><br>• For binary responder variables logistic regression models will be fitted to examine the relationship between treatment group and the endpoint of interest. No covariates will be included in these analyses.<br><br>The LS means for the treatment groups and the LS mean differences from placebo will be estimated and back-transformed to provide the odds and odds ratios, respectively.<br><br>A results table will present the odds of response for each treatment group (and 95% CIs), as well as the odds ratio of each active treatment group to placebo and the odds ratio of 90 mg BID to 180<br><br>The odds ratio will be plotted along with the corresponding 95% CIs, with a reference line at y=1.<br><br>• For Time to Event data Kaplan-Meier estimates and plots will be presented for each treatment group. Treatment groups will only be included in this analysis if there are at least 8 participants in the group, within the population of interest. |
| **Subgroup Analyses** |
| Please see the table above under model specification for list of covariates that are included in each model. |

## 7.3. Exploratory Efficacy Analyses

### 7.3.1. Endpoint / Variables

All exploratory efficacy endpoints are listed under Section 2.2.

### 7.3.2. Summary Measure

The exploratory efficacy endpoints will be summarised as the primary and secondary efficacy endpoints.

**Responders Time to Worsening:**

Responders Time to Worsening of Itch (in sub-population of Responders (>=2 point improvement from baseline)). The time to worsening of itch after Week 16 (when on single-blind placebo) in those who have responded at Week 16 (Responders (>=2 point improvement) population).

**PRO: Other eDiary Itch questions:**

In addition to the primary endpoint further questions in the eDiary capture information on the itch experienced by participants. Overall itch, sleep and fatigue use the same NRS (0-10), while the others are categorical.

Mean Overall Itch Score

Other daily eDiary questions

**Other PRO measurements:**

EQ-5D-5L (For the categories of EQ-5D-5L)

EQ-5D-5L (Separately for the derived EQ-5D-5L Utility Index Score see Section 12.6.3 for more details and the EQ-5D-5L VAS)

PGI-S questions

PGI-C questions

BDI-II (total BDI-II score)

PTEA

**Actigraphy data:**

- Mean duration of scratching Events
- Mean number of scratching events
- Sleep Information

### 7.3.3. Population of Interest

The exploratory efficacy analyses will be based on the ITT population, unless otherwise specified.

### 7.3.4. Statistical Analyses / Methods

- Details of the planned displays are provided in Appendix 12: List of Data Displays and will be based on GSK data standards and statistical principles.

- Unless otherwise specified, endpoints / variables defined in Section 7.3.1 will be summarised using descriptive statistics, graphically presented (where appropriate) and listed.

  **Actigraphy:** Actigraphy data will only be analysed if there is sufficient reliable data for the analysis to be meaningful for the subset of the participants recruited to the study who take part in an actigraphy study in which they will wear monitors to evaluate their movement. Algorithms will then transform this information into measures of sleep efficiency and number and duration of scratching events during their rest period. The monitors will be worn during at least 5 nights over a 7 day period, on 3 separate occasions throughout the study and for each night the mean

number of scratching events and mean duration of scratching events per hour, during the rest period, will be provided.

Actigraphy data will be listed only.

# 8. SAFETY ANALYSES

The safety analyses will be based on the Safety population, unless otherwise specified.

## 8.1. Adverse Events Analyses

Adverse events analyses including the analysis of adverse events (AEs), Serious (SAEs) and other significant AEs will be based on GSK Core Data Standards. The details of the planned displays are provided in Appendix 12: List of Data Displays.

Additional analysis will be provided for abdominal pain AEs in addition to the standard analyses as describe in Appendix 12. The Kaplan-Meier curve will be provided for time to event analyses. Moreover, the duration of abdominal pain will be assessed using swimmer plots. The number of participants who experienced abdominal pain will be provided using summary statistics.

Please refer to Section 12.6.4 to identify abdominal pain based on MedDRA terms.

## 8.2. Adverse Events of Special Interest Analyses

A comprehensive list of MedDRA terms based on clinical review will be used to identify each type of event. Due to the special interest in GI events, time to event analysis will be provided to assess both the duration and offset of GI event. The Kaplan-Meier curve will be provided for time to event analyses. Moreover, the duration of diarrhoea and will be assessed using swimmer plots. The number of participants who experienced diarrhoea will be provided; in addition, the duration of both GI events in the same participants will be produced using summary statistics.

Please refer to Section 12.6.4 to identify diarrhoea based on MedDRA terms.

## 8.3. Clinical Laboratory Analyses

Laboratory evaluations including the analyses of Chemistry laboratory tests, Haematology laboratory tests, Urinalysis, and liver function tests will be based on GSK Core Data Standards. The details of the planned displays are in Appendix 12: List of Data Displays.

## 8.4. Other Safety Analyses

The analyses of non-laboratory safety test results including ECGs and vital signs will be based on GSK Core Data Standards, unless otherwise specified. The details of the planned displays are presented in Appendix 12: List of Data Displays.

There are two sets of questions in the eDiary which will also be summarised within the safety analysis: the Gastrointestinal Symptoms Rating Scale (GSRS) questionnaire and specific questions pertaining to the perceived gastrointestinal (GI) side effects of the study treatment, namely diarrhoea and abdominal pain, these will be presented too.

The exploratory safety endpoints as listed under Section 2.2. will be summarised using descriptive statistics and will be listed as appropriate. Details of the planned displays are provided in Appendix 12: List of Data Displays and will be based on GSK data standards and statistical principles.

Mean change from baseline at week 16 in Vitamins A, D, E, K and FOBT will be summarised and listed by treatment group.

Prior to 2019 FOBT card (InsureFIT) required 2 samples of stool/fecal and were collected by participant at home prior to the visit. In 2018, the FOBT card (InsureFIT) was discontinued by the manufacturer. In early 2019, the replacement FOBT card (InsureONE) was then used for the remainder of the study.
The new FOBT InsureONE card only requires **1** stool/fecal sample to be collected at home, prior to the visit.

# 9. PHARMACOKINETIC ANALYSES

## 9.1. Primary Pharmacokinetic Analyses

### 9.1.1. Endpoint / Variables

#### 9.1.1.1. Drug Concentration Measures and Data Handling

Refer to Appendix 5: Data Display Standards & Handling Conventions (Section 12.5.3) Reporting Standards for Pharmacokinetic) and see below

If a linerixibat concentration is reported as missing in the linerixibat treatment groups, then this will be reported as missing, if a concentration at V4, V5 or V6 is reported as NQ, then this will be handled as ½ LLQ (or ½ of 10pg/mL=5pg/mL) for inclusion in plots and PK parameter summary tables.

If a UDCA, GUDCA or TUDCA concentration is missing or NQ, then these will be reported and handled as missing or NQ, respectively. The NQ will be handled as missing for plots and PK parameter summary tables. Participants on placebo are included and will have UDCA, GUDCA and TUDCA conjugates plasma concentrations.

#### 9.1.1.2. Pharmacokinetic Concentrations for linerixibat, UDCA, GUDCA and TUDCA

Pharmacokinetic parameters are not calculable based on the sampling scheme and average PK concentrations will be used. All calculations of PK concentrations will be based on actual sampling times that fall within the set PK windows described below (±0.75 hour of the PK window). Population PK and/or PD analysis may be conducted and will be described in a separate report.

Use the following guide for PK concentrations handling:

- If a $C_{ave,V}$ has a missing value for or both $C_{v,1\text{-}3h}$ or $C_{v,5\text{-}8h}$, then don't report $C_{ave,V}$
- For BID, the PK concentrations are associated with the first 8 hours of the first dosing interval of the day
- For QD, the PK concentrations are associated with the first 8 hours of the daily dose

| PK Parameter | Description |
|---|---|
| $C_{V3,1\text{-}3h}$ | Average concentration of PK analyte in time window 1-3 hr at visit 3, week 4 |
| $C_{V3,\ 5\text{-}8h}$ | Average concentration of PK analyte in time window 5-8 hr at visit 3, week 4 |
| $Cave_{,V3}$ | Average concentration in first 8 hr post-dose at visit 3, week 4 (i.e., = ($C_{V3,1\text{-}3h}$ + $C_{V3,\ 5\text{-}8h}$)/2) |
| $C_{V4,1\text{-}3h}$ | Average concentration of PK analyte in time window 1-3 hr at visit 4, week 8 |

| PK Parameter | Description |
|---|---|
| $C_{V4,5\text{-}8h}$ | Average concentration of PK analyte in time window 5-8 hr at visit 4, week 8 |
| $Cave_{,V4}$ | Average concentration in first 8 hr post-dose at visit 4, week 8 (i.e., = $(C_{V4,1\text{-}3h} + C_{V4,\ 5\text{-}8h})/2$) |
| $C_{V5,1\text{-}3h}$ | Average concentration of PK analyte in time window 1-3 hr at visit 5, week 12 |
| $C_{V5,5\text{-}8h}$ | Average concentration of PK analyte in time window 5-8 hr at visit 5, week 12 |
| $Cave_{,V5}$ | Average concentration in first 8 hr post-dose at visit 5, week 12 (i.e., = $(C_{V5,1\text{-}3h} + C_{V5,\ 5\text{-}8h})/2$) |
| $C_{V6,1\text{-}3h}$ | Average concentration of PK analyte in time window 1-3 hr at visit 6, week 16 |
| $C_{V6,5\text{-}8h}$ | Average concentration of PK analyte in time window 5-8 hr at visit 6, week 16 |
| $Cave_{,V6}$ | Average concentration in first 8 hr post-dose at visit 6, week 16 (i.e., = $(C_{V6,1\text{-}3h} + C_{V6,\ 5\text{-}8h})/2$) |
| Cmax | Highest concentration reported at any time in the 16 weeks |

For BID, the PK concentrations are associated with the first 8 hours of the first dosing interval of the day
For QD, the PK concentrations are associated with the first 8 hours of the daily dose
If a Cave,V has a missing value for or both Cv,1-3h or Cv,5-8h, then don't report Cave,V

### 9.1.2. Summary Measure

For each PK window listed in section above: Arithmetic mean, standard deviation, CV%a, 95% CI, median, range, geometric mean, CV%g, and 95% CI for geometric mean.

### 9.1.3. Population of Interest

The primary pharmacokinetic analyses will be based on the Pharmacokinetic population, unless otherwise specified.

### 9.1.4. Statistical Analyses / Methods

Details of the planned displays are provided in Section 12.12.9 and Section 12.12.10, List of Data Displays and will be based on GSK Data Standards and statistical principles.

Unless otherwise specified, endpoints / variables defined in Section 9.1.1 will be summarised using descriptive statistics, graphically presented (where appropriate) and listed.

The PK analysis will be based on appropriate sub-groups of the PK population. The analysis will consist of summaries of plasma concentrations, by treatment group, visit and time point as well as profile plots for each subject, where appropriate.

Linerixibat will only be summarised and presented for subjects who receive active treatment (20 mg QD, 90 mg QD, 180 mg QD and 90 mg BID) and from whom samples are available.

UDCA will only be summarised and presented for subjects who are taking UDCA (and for whom UDCA samples are available).

Scatter plots of concentration and actual sampling time data for linerixibat, UDCA, GUDCA and TUDCA will be produced. Plots will separate groups by linerixibat treatment alone and linerixibat + UDCA co-treatment groups.

Line plots of change from baseline in eGFR at V4 and V5 vs linerixibat Cave plasma concentrations at V4 and V5, respectively, will be produced. (renal impairment exploration) Box plots of linerixibat, UDCA, GUDCA and TUDCA average concentrations vs treatment groups will be produced (dose response and DDI exploration)

# 10. PHARMACODYNAMIC BIOMARKER ANALYSES

## 10.1. Primary/Secondary/Exploratory Pharmacodynamic Biomarker Analyses

Refer to Appendix 5: Data Display Standards & Handling Conventions (Section 12.5.3 Reporting Standards for Pharmacodynamics) and see below

If a PD biomarker concentration is missing it will be reported and handled as missing. If a PD biomarker is BQL it will be handled as (LLQ/2). If a PD biomarker is above the quantitation limit (AQL) will be handled according to existing data conventions.

Note: For ELF test, there is no LLQ due to nature of measurement.

### 10.1.1. Endpoint / Variables

**List of Biomarkers of PBC and Bile Acid Physiology:**

- Total Serum Bile Acids (TSBA, available as sum of conjugated and unconjugated, not available separated)
- Individual serum bile acid species (measurement provided from laboratory as sum of unconjugated + conjugated) for: cholic acid (CA), chenodeoxycholic acid (CDCA), deoxycholic acid (DCA)
  - (e.g., CA = CA + glycocholic acid (GCA) + taurocholic acid (TCA), or CDCA = CDCA + GCDCA + TCDCA, or DCA = DCA + GDCA + TDCA)
- Serum C4 (7α-hydroxy-4-cholesten-3-one)
- Serum autotaxin enzyme protein concentrations (not autotaxin enzyme activity)
- Serum Fibroblast Growth Factor –19 (FGF-19)
- Enhanced Liver Fibrosis (ELF) Test
- Alkaline Phosphatase (ALP)
- Total bilirubin

### 10.1.2. Summary Measure

Summary measure will include change from baseline and absolute values for each visit and each dose group for biomarkers.

The absolute values and change from baseline for all biomarkers listed in 10.1.1 will be summarized and tabulated and plotted by treatment and visit

The absolute values and change from baseline for lipid biomarkers will be summarized and tabulated by treatment and visit.

### 10.1.3. Population of Interest

The PD biomarker analyses will be based on ITT.

### 10.1.4. Statistical Analyses / Methods

The PD analysis will concentrate on the biomarkers of PBC and bile acid physiology as outlined in Section 10.1.1. Details of the planned displays are provided in Appendix 12.12.10: List of Data Displays and will be based on GSK Data Standards and statistical principles.

**PD Biomarker Correlation Figures and Tables**

Details of the planned displays are provided in Section 12.12.11 and Section 12.12.12, List of Data Displays and will be based on GSK Data Standards and statistical principles.

Unless otherwise specified, endpoints / variables defined in Section 10.1.1 will be summarised using descriptive statistics, graphically presented (where appropriate) and listed

Of note, C4, FGF19, ELF are not pruritogens and not included in correlations section.

Correlation analyses of mean changes from baseline for ALP and mean changes of baseline for TSBA, CA, CDCA, DCA, UDCA, GUDCA and TUDCA for V4, V5 and V6 (Section 12.12.11)

Absolute and change from baseline in ALP for subject taking UDCA by visit and treatment group.

Correlation analyses will be produced between baseline of PD biomarkers (see list from 10.1.1) at visit 6 with baseline NRS itch scores by treatment groups.

Correlation analyses will be produced between change from baseline of PD biomarkers (as listed form 10.1.1) at visit 6 with change from baseline NRS itch scores at visit 6 by treatment groups.

Pearson correlation will be produced given that the data follow the normal distribution, normality test such as Kologorov-Smirov will be used. However, if the normality assumption failed a log transformation will be applied and if the normality assumption failed after the transformation a Spearman correlation will be produced.

Unless otherwise specified, endpoints / variables defined in Section 2.2. will be summarised using descriptive statistics, graphically presented (where appropriate) and listed.

#### 10.1.4.1. Statistical Methodology Specification

| **Endpoint / Variables** |
|---|
| Endpoints are listed 10.1.1 |
| **Model Specification** |
| ANCOVA models for change from baseline will be fitted adjusting for baseline and treatment groups.<br>Listings, summary statistics, and line plots to explore time course and dose response will be produced.<br>PD biomarkers correlations as stated in 10.1.1 and as described in 12.12.11 and 12.12.12. |
| **Model Results Presentation** |
| Mean difference for each treatment group from placebo, along with mean difference between 90 mg BID and 180 mg QD will be presented from the ANCOVA model. |

# 11. REFERENCES

GlaxoSmithKline Document Number 2010N111289_06 Investigator's Brochure. Linerixibat (GSK2330672) Investigator's Brochure. Date 29-AUG-2019

GlaxoSmithKline Document Number 2016N280566_01 Study ID 201000. A randomised, double-blind, multi-dose, placebo-controlled study to evaluate the efficacy, safety and tolerability of GSK2330672 administration for the treatment of pruritus in patients with primary biliary cholangitis. (GLIMMER: GSK2330672 triaL of Ibat inhibition with Multidose Measurement for Evaluation of Response).Report Date 04-Nov-2017.

GlaxoSmithKline Document Number N2015N268599_01 Clinical Pharmacology Study Report. A randomised, double-blind, placebo-controlled study to evaluate the safety, tolerability, pharmacokinetics and pharmacodynamics of repeat doses of GSK2330672 administration in patients with primary biliary cirrhosis (PBS) and symptoms of pruritus. Date 15-DEC-2016

Hegade, VS., Kendrick SFW., Dobbins, RL., et al. Effect of ileal bile acid transporter inhibitor GSK2330672 on pruritus in primary biliary cholangitis: a double-blind, randomised, placebo-controlled, crossover, phase 2a study. The Lancet, 389:1114, March 18, 2017.

Kirby, S., Brain, P., Jones, B. Fitting Emax models to clinical trial dose-response data. Pharmaceut. Statist. 2011; 10: 143-149.

van Hout B, Janssen MF, et al. Interim scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health 2012 Jul-Aug;15(5):708-15.

# 12. APPENDICES

## 12.1. Appendix 1: Protocol Deviation Management and Definitions for Per Protocol Population

### 12.1.1. Exclusions from Per Protocol Population

A participant meeting any of the following criteria will be excluded from the Per Protocol population:

| | Exclusion Description |
|---|---|
| **Protocol Deviations** | |
| | Mean Worst Daily Itch not ≥ 3 for at least 5 of the 7 days prior to the Baseline visit |
| | Other Important Protocol Deviations (as defined in Protocol Deviation Management Plan (PDMP) and discussed at blinded data review. |
| **Treatment Compliance** | |
| | Study Treatment Compliance<br>Less than 80% or Greater than 120% total main treatment period (double blind treatment) compliance in MSP according to tablets taken (V3 to V6) |
| | Study Treatment Compliance<br>According to the eDiary, did not take study treatment on 3 or more mornings in the 7 days prior to V6. |
| | Study Treatment Compliance<br>According to the eDiary, did not take study treatment on 3 or more evenings in the 7 days prior to V6. |
| **Endpoint Compliance** | |
| | End-Point Compliance at V6<br>Worst Daily Itch is missing for 3 or more of the 7 days prior to V6 |
| **Visit Timing** | |
| | Missed V6 |
| | V6 outside of Analysis Time Window (where Analysis Time Window is defined as Target day ± 7 days, i.e. Day 105 - 119) |
| **Medications** | |
| | **Prohibited Medication use**<br>**Other IBAT inhibitors:** Use of any other IBAT inhibitor is not permitted at any time during the study.<br>**Obeticholic acid:** Use is not permitted at any time during the study and obeticholic acid should be discontinued at least 8 weeks before the start of the Initial Study Period. |

| | Exclusion Description |
|---|---|
| | **Cholestyramine, colesevelam, colestipol or colestimide:** Use of any of these agents is not permitted during the Initial or Main Study Periods and will be discontinued no later than one full day (i.e., no later than Day -2) before the start of the Initial Study Period (Visit 2). During the Final Study Period and the Follow Up Period, use of these agents may be permitted. |
| | **Permitted Medication use**<br>Please see the Protocol Section 7.7.2. for details and list of permitted medication use. |

## 12.2. Appendix 2: Schedule of Activities

### 12.2.1. Protocol Defined Schedule of Events

**Schedule of Activities Table**

| | Screening | Initial Study Period | Main Study Period | | | | Final Study Period | Follow-up Period (FU) | Early End of Treatment or Study Withdrawal Assessments[1] |
|---|---|---|---|---|---|---|---|---|---|
| Visit Name | **Visit 1 (Screening)** | **Visit 2 (Day 1)** | **Visit 3[11] (Baseline/ Week 4)** | **Visit 4 (Week 8)** | **Visit 5 (Week 12)** | **Visit 6 (Week 16)** | **Visit 7 (Week 20)** | **Visit 8 (End of FU Phone Visit/ Week 24)** | |
| Day number (window) | Day -45 to -1 | Day 1 | Day 28 (Day 28 to 35) | Day 56 (Day 56 to 63) | Day 84 (Day 84 to 91) | Day 112 (Day 112 to 119) | Day 140 (Day 140 to 147) | Day 168 (Day 168 to 175) | |
| **Procedure** | | | | | | | | | |
| Screening procedures | X[2] | | | | | | | | |
| Brief physical exam inc. weight | | X | X | X | X | X | X | | X |
| Concomitant medications | X[2] | X | X | X | X | X | X | | X |
| 12-lead ECG | X[2] | X | X | X | X | X | X | | X |
| Vital signs (HR and BP) | X[2] | X | X | X | X | X | X | | X |
| Pregnancy test (urine dipstick) in WOCBP | X[2] | X[3] | X | X | X | X | X | | X |
| Clinical laboratory tests[4] | X[2] | X | X | X | X | X | X | | X |
| AE Assessment | | X | X | X | X | X | X | | X |
| Study treatment dispensing | | X | X | X | X | X | | | |
| Study treatment administration | | <----------Participants administer morning and evening each day----------> | | | | | | | |
| Study treatment compliance | | | X | X | X | X | X | | |

**CONFIDENTIAL**



| | **Screening** | **Initial Study Period** | **Main Study Period** | | | | **Final Study Period** | **Follow-up Period (FU)** | **Early End of Treatment or Study Withdrawal Assessments[1]** |
|---|---|---|---|---|---|---|---|---|---|
| Visit Name | **Visit 1 (Screening)** | **Visit 2 (Day 1)** | **Visit 3[11] (Baseline/ Week 4)** | **Visit 4 (Week 8)** | **Visit 5 (Week 12)** | **Visit 6 (Week 16)** | **Visit 7 (Week 20)** | **Visit 8 (End of FU Phone Visit/ Week 24)** | |
| Day number (window) | Day -45 to -1 | Day 1 | Day 28 (Day 28 to 35) | Day 56 (Day 56 to 63) | Day 84 (Day 84 to 91) | Day 112 (Day 112 to 119) | Day 140 (Day 140 to 147) | Day 168 (Day 168 to 175) | |
| Final eligibility and Randomization | | | X[12] | | | | | | |
| PD/Biomarker blood samples[4] | | | X | X | X | X | X | | X |
| FOBT[5] | | | X | | | X | | | |
| Genetics sample[6] | | | X | | | | | | |
| PK samples[7] | | | X | X | X | | | | |
| Symptom eDiary | | <----------Morning and evening each day----------> | | | | | | | |
| GSRS and GI symptoms on eDiary | | <----------Once weekly----------> | | | | | | | |
| PBC-40 Scale | X[2] | X | X | X | X | X | X | | X |
| 5D-Itch Scale | | X | X | | | X | X | | X |
| EuroQOL 5D-5L | | X | X | | | X | | | X |
| PGI-S | | X | X | X | X | X | X | | X |
| PGI-C | | | X | X | X | X | X | | X |
| BDI-II | X[2] | | X | | | X | | | X |
| Actigraphy[8] | | X | | X | X | | | | |
| Participant Treatment Experience Assessment | | | | | | X | | | X[9] |
| Telephone review[10] | | | | | | | | X | |

Footnotes:

1. Early End of Treatment Assessments to be completed for participants who prematurely discontinue study treatment following randomization (see Section 8.1 of the protocol). Study Withdrawal Assessments to be completed for participants who withdraw consent for any further participation in the study following randomization (see Section 8.2 of the protocol). Note: If a participant discontinues study treatment at the same time they withdraw from the study, only the End of Treatment Assessments are required.
2. See the table below for detailed description of Screening Procedures.
3. Urine dipstick pregnancy test will be standard unless serum testing is required by local regulation or IRB/IEC. Urine dipstick test at Visit 2 not required if the screening test was performed <1 week of Visit 2.
4. For clinical laboratory tests see Section 9.4.4. of the protocol and biomarkers see Section 9.8 of the protocol. All laboratory and biomarker assessments will be fasting (water, study treatment and other medications are permitted) except for Screening Visit.
5. FOBT (card test on 2 different stool/fecal samples) will be performed by participant at home prior to the visit.
6. Genetics sample in randomised participants only, to be collected preferably at Visit 3, but may be collected at Visit 4.
7. PK sample at Visit 3 for participants on UDCA only. At other visits, two plasma samples will be collected on each PK occasion in all participants (see Section 9.5 of the protocol). If PK sample collection is not collected at Visit 5 it may alternatively be collected at Visit 6. Note: PK sample collection is not required for participants who have prematurely discontinued study treatment and have not taken study treatment on the 3 or more days prior to the Visit.
8. Participants who give consent for the Actigraphy Sub-study will be given actigraphy monitors at Visit 2, Visit 4 and Visit 5 with activity measurements performed over at least 5 nights during Weeks 2 or 3, 9 or 10, and 13 or 14 respectively (see Section 9.1.3 of the protocol).
9. Participant Treatment Experience Assessment only required for Early End of Treatment Assessments.
10. End of Follow Up Telephone Visit to review participant's itch and anti-pruritus medications.
11. In the event of a **confirmed** baseline (run-in) failure at Visit 3, blood, urine, ECG, and TrialMax Slate assessments are not required, and no further study treatment will be dispensed. If the FOBT has been collected and returned by the participant, it should be sent to the laboratory for evaluation. Refer to Section 6.5.2 of the protocol for more information.
12. Randomization criteria must be met at this visit for a participant to be randomised. Refer to Section 6.3 of the protocol.

Abbreviations:
AE= Adverse Event; BDI-II= Beck Depression Inventory-II; BP= Blood Pressure; ECG= Electrocardiogram; EQ-5D= EuroQOL 5D-5L; FOBT= Fecal Occult Blood Test; GI= gastrointestinal; GSRS= Gastrointestinal Symptoms Rating Scale; HR= Heart Rate; IEC= Independent Ethics Committee; IRB= Institutional Review Board; NRS= Numerical Rating Scale; PD= Pharmacodynamic; PGI-S= Patient Global Impression of Severity; PGI-C= Patient Global Impression of Change; PK= Pharmacokinetics; UDCA= Ursodeoxycholic acid; WOCBP= Women of Child Bearing Potential

- For assessments scheduled at the same visit, where possible they should be performed in the following order: fasting blood tests and pharmacokinetic (PK) sample (1-3 hour post-dose and 5-8 hours), patient reported outcomes (PROs) [i.e., PGI-S, PGI-C, PBC-40, 5D-Itch, EQ-5D, BDI-II, Participant Treatment Experience Assessment], ECGs, vital signs, physical exam, other assessments, PK sample (5-8 hours post-dose). See Section 9 of the protocol for further details.
- All reasonable attempts should be made to ensure compliance with the visit schedule, however participants will not be considered out of compliance if visit windows extend because of extraordinary events that make it impossible for participants to complete a visit within the required window. Determination of the maximum visit window deviation is at the discretion of the Medical Monitor.
- The timing and number of planned study assessments specifically PK or pharmacodynamic (PD)/biomarkers assessments may be altered during the course of the study based on newly available data (e.g., to obtain data closer to the time of peak plasma concentrations) to ensure appropriate monitoring.
- Any changes in the timing or addition of time points for any planned study assessments specified above must be documented and approved by the relevant study team member and then archived in the sponsor and site study files, but will not constitute a protocol amendment. The Institutional Review Boards (IRB)/ Independent Ethics Committees (IEC) will be informed of any safety issues that require alteration of the safety monitoring scheme or amendment of the informed consent form (ICF).

**Screening Procedures Table**

| | Screening Visit 1 | Notes |
|---|---|---|
| Window | Day -45 to -1 | |
| **Procedure** | | |
| Informed Consent | X | |
| Inclusion/Exclusion Criteria | X | |
| Demography | X | |
| Full physical examination including height and weight | X | |
| Medical history (includes substance usage, family history of liver disease, cancer and cardiovascular disease) | X | Substances: alcohol and tobacco |
| Past and current medical conditions, including PBC and pruritus history, and associated medications | X | Includes use and response to UDCA, cholestyramine, colesevalem and other anti-pruritus medications |
| PBC-40 Scale | X | |
| BDI-II | X | |
| 12-lead ECG | X | |
| Vital signs | X | |
| Laboratory assessments (non-fasting), including liver chemistries | X | |
| Urine dipstick pregnancy test (WOCBP only)[1] | X | |
| Hepatitis B and C screening | X | |

1. Footnotes:
1. Urine testing will be standard unless serum testing is required by local regulation or IRB/IEC.

Abbreviations
BDI-II= Beck Depression Inventory-II
ECG= electrocardiogram
PBC= Primary Biliary Cholangitis
WOCBP = Women Of Child Bearing Potential
UDCA = Ursodeoxycholic acid

## 12.3. Appendix 3: Assessment Windows

### 12.3.1. Definitions of Assessment Windows for Analyses

| Analysis Time point | End of Week | Target | Protocol Window | | Analysis Window | |
|---|---|---|---|---|---|---|
| | | | Beginning Time point | Ending Time point | Beginning Time point | Ending Time point |
| Visit 1 | NA | Screening | Day -45 | Day -1 | Day -45 | Day -1 |
| Visit 2 | NA | Day 1 | Day 1 | Day 1 | Day 1 | Day 1 |
| Visit 3 (Baseline) | Week 4 | Day 28 | Day 28 | Day 35 | Day 21 | Day 35 |
| Visit 4 | Week 8 | Day 56 | Day 56 | Day 63 | Day 49 | Day 63 |
| Visit 5 | Week 12 | Day 84 | Day 84 | Day 91 | Day 77 | Day 91 |
| Visit 6 | Week 16 | Day 112 | Day 112 | Day 119 | Day 105 | Day 119 |
| Visit 7 | Week 20 | Day 140 | Day 140 | Day 147 | Day 133 | Day 147 |
| Visit 8 | Week 24 | Day 168 | Day 168 | Day 175 | Day 161 | Day 175 |

NA=Not Applicable.

Note: Visits outside the protocol windows will be recorded as protocol deviations. For the purposes of the analysis visits will only be excluded from the PP figures, summaries, and statistical analyses if they fall outside the analysis windows.

**Definitions of Days for Diary Assessments**

| Analysis Time point | Day of Analysis Time point | Labels for Summaries | Diary Data Days used in Summaries |
|---|---|---|---|
| Visit 1 | V1 | Screening (V1) | NA |
| Visit 2 | V2 | Day 1 (V2) | NA |
| Visit 3 (Baseline) | V3 | Week 4 (V3) | V3-7 to V3-1 |
| Visit 4 | V4 | Week 8 (V4) | V4-7 to V4-1 |
| Visit 5 | V5 | Week 12 (V5) | V5-7 to V5-1 |
| Visit 6 | V6 | Week 16 (V6) | V6-7 to V6-1 |
| Visit 7 | V7 | Week 20 (V7) | V7-7 to V7-1 |
| Visit 8 | V8 | Week 24 (V8) | NA |

Vx = Day of Visit x.
FU = Follow-up.
NA=Not Applicable.

## 12.4. Appendix 4: Study Phases and Treatment Emergent Adverse Events

### 12.4.1. Study Phases

Assessments and events will be classified according to the time of occurrence relative to the ISP, MSP, FSP and FU

| **Treatment Period** | Abbreviation | **Treatment Details** | **Definition** | **Notes** |
|---|---|---|---|---|
| Screening | | None | Date < V2 | |
| Initial Study | ISP | placebo only | V2 ≤ Date < V3 | |
| Main Study | MSP | Randomised | V3 ≤ Date ≤ V6 | Responder Days will use data from V3+1 to V6-1 |
| Final Study | FSP | placebo only | V6 < Date < V7 | Time to Worsening will use data from V6+1 to V7-1 |
| Follow-up | FU | None | V7 ≤ Date ≤ V8 | |

Vx = day of Visit x (Section 14.3.1.)

Assessments and events will be classified according to time of occurrence relative to the start and/or stop date of the main study period.

#### 12.4.1.1. Treatment Phase All Data

| **Study Phase** | **Definition** |
|---|---|
| Pre-Treatment | Date < Date of first Randomised Treatment |
| On-Treatment | Date of first Randomised Treatment ≤ Date ≤ Date of last Randomised Treatment |
| Post-Treatment | Date of last Randomised Treatment < Date |

- **Date** = Assessment/Start Date

#### 12.4.1.2. Treatment Phase for AE Data

| **Treatment State** | **Definition** |
|---|---|
| Pre-Treatment | AE Start Date < Date of first Randomised Treatment |
| On-Treatment | Date of first Randomised Treatment ≤ AE Start Date ≤ Date of last Randomised Treatment + 2 |
| Post-Treatment | Date of last Randomised Treatment + 2 < AE Starts Date |
| Onset Time Since 1st Dose (MSP Day) | If Date of first Randomised Treatment > AE Onset Date = AE Onset Date – Date of first Randomised Treatment<br>If Date of first Randomised Treatment ≤ AE Onset Date = AE Onset Date – Date of first Randomised Treatment +1<br>Missing otherwise. |
| Duration (Days) | AE Resolution Date – AE Onset Date + 1 |
| Drug-related | If relationship is marked 'YES' on eCRF |

### 12.4.1.3. Study Phases for Concomitant Medication

| Study Phase | Definition |
|---|---|
| Prior | Any medication or vaccine (including over-the-counter or prescription medicines, topical agents including topical corticosteroids, vitamins, and/or herbal supplements) that the participant has taken in the 3 months prior to Screening. |
| Concomitant | Any medication receiving at the time of enrolment or receives during the study period (Initial Study Period, Main Study Period, and Final Study Period) A medication could potentially be both prior and concomitant. |

**NOTES:**

- Please refer to Appendix 7: Reporting Standards for Missing Data for handling of missing and partial dates for concomitant medication. Use the rules in this table if concomitant medication date is completely missing.

### 12.4.2. Treatment Emergent Flag for Adverse Events

| Flag | Definition |
|---|---|
| Treatment Emergent | • If AE onset date is on or after treatment (Main study period or Final study period) start date & on or before treatment stop date. (plus Final study period or protocol-specified time limit (e.g. half-life of drug, certain number of days, etc.).<br>• Study Treatment Start Date ≤ AE Start Date ≤ Study Treatment Stop Date [+ 2 days].<br>• For studies with greater than one treatment period (e.g., crossover study), if AE onset is during one period and worsens during a later period it would be counted in both periods. For the initial period the logic would be as above. For the later period the logic would use the treatment dates associated with the later period:<br>• Treatment Period Start Date ≤ AE Worsening Date ≤ Study Treatment Stop Date [+ 2 days]. |

**NOTES:**

- If the study treatment stop date is missing, then the AE will be considered to be On-Treatment.
- Time of study treatment dosing and start/stop time of AEs should be considered, if collected.

## 12.5. Appendix 5: Data Display Standards & Handling Conventions

### 12.5.1. Reporting Process

| Software | |
|---|---|
| • The currently supported versions of SAS software will be used. | |
| **Reporting Area** | |
| HARP Server | : us1salx00259 |
| HARP Compound | : \ARPROD\ GSK2330672\mid201000 |
| **Analysis Datasets** | |
| • Analysis datasets will be created according to CDISC standards (SDTM IG Version 3.2 & ADaM IG Version 2.0).<br>• For creation of ADaM datasets (ADCM/ADAE), the same version of dictionary datasets will be implemented for conversion from SI to SDTM. | |
| **Generation of RTF Files** | |
| • RTF files will be generated for all the tables described in the RAP. | |

### 12.5.2. Reporting Standards

| General |
|---|
| • The current GSK Integrated Data Standards Library (IDSL) will be applied for reporting, unless otherwise stated (IDSL Standards Location: https://spope.gsk.com/sites/IDSLLibrary/SitePages/Home.aspx):<br>  • 4.03 to 4.23: General Principles<br>  • 5.01 to 5.08: Principles Related to Data Listings<br>  • 6.01 to 6.11: Principles Related to Summary Tables<br>  • 7.01 to 7.13: Principles Related to Graphics<br>• Do not include participant level listings in the main body of the GSK Clinical Study Report. All participant level listings should be located in the modular appendices as ICH or non-ICH listings |
| **Formats** |
| • GSK IDSL Statistical Principles (5.03 & 6.06.3) for decimal places (DP's) will be adopted for reporting of data based on the raw data collected, unless otherwise stated.<br>• Numeric data will be reported at the precision collected on the eCRF.<br>• The reported precision from non eCRF sources will follow the IDSL statistical principles but may be adjusted to a clinically interpretable number of DP's.<br>  • For [Insert Endpoint / Parameter] the following DP's places will be applied:<br>  • Summary Statistics:<br>  • Listings: |
| **Planned and Actual Time** |
| • Reporting for tables, figures and formal statistical analyses:<br>  • Planned time relative to dosing will be used in figures, summaries, statistical analyses and calculation of any derived parameters, unless otherwise stated. |

- The impact of any major deviation from the planned assessment times and/or scheduled visit days on the analyses and interpretation of the results will be assessed as appropriate.

- Reporting for Data Listings:
  - Planned and actual time relative to study drug dosing will be shown in listings (Refer to IDSL Statistical Principle 5.05.1).
  - Unscheduled or unplanned readings will be presented within the participant's listings.

**Unscheduled Visits**

- Unscheduled visits will not be included in summary tables except where specifically justified and flagged at the blind data review meeting. Such data (decided on an individual case basis) will replace the respective nominal visit in summary statistics.
- In most cases, unscheduled visits will not contribute to summary statistics or plots of summary statistics or box plots.

- Unscheduled visits will be included in figures of individual data where profile plots are presented over time and will be presented within any relevant listings, and in tables of values of potential clinical importance.

- All unscheduled visits will be included in listings in chronological order with respect to the scheduled visits.

**Descriptive Summary Statistics**

| | |
|---|---|
| Continuous Data | Refer to IDSL Statistical Principle 6.06.1 |
| Categorical Data | N, n, frequency, % |

**Graphical Displays**

- Refer to IDSL Statistical Principals 7.01 to 7.13.

### 12.5.3. Reporting Standards for Pharmacokinetic and Pharmacodynamic

Refer to 9.1.1.1 for PK concentration and 9.1.1.2 for PK concentration data handling and reporting.

Refer to 10.1 for PD concentration data handling and reporting

| **Pharmacokinetic Concentration Data** | |
|---|---|
| | Refer to 9.1.1.1 for PK concentration<br>Note: Concentration values for BQL will be imputed using LLQ/2 as per GUI_51487 |
| Descriptive Summary Statistics, Graphical Displays and Listings | Refer to IDSL PK Display Standards.<br>Refer to IDSL Statistical Principle 6.06.1.<br>Note: Concentration values for BQL will be imputed using LLQ/2 as per GUI_51487 for descriptive summary statistics/analysis and summarized graphical displays only. |
| **Pharmacokinetic Parameter Derivation** | |
| PK Parameter to be Derived by Programmer | NA |
| **Pharmacokinetic Parameter Data** | |
| Is NQ impacted PK Parameters Rule Being Followed | NA |
| Descriptive Summary Statistics, Graphical Displays and Listings | Refer to IDSL PK Display Standards.<br>and Section 9.1.1.1 and Section 9.1.1.2 |

## 12.6. Appendix 6: Derived and Transformed Data

### 12.6.1. General

| Multiple Measurements at One Analysis Time Point |
|---|
| If there is more than one measurement at any given nominal time point, the mean of all the measurements will be calculated and used in any derivation of summary statistics, but if listed, all measurements will be presented.<br><br>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.<br><br>If there is more than one entry in the e-diary data for either 'Worst Itch' or 'Sleep' on any given day and timepoint (AM or PM) then it will use the worst score for that day or timepoint. The derivation of mean worst daily itch and mean sleep score will continue as in Sections 12.6.3, respectively, if there are enough non-missing data to satisfy the conditions of the derivation.<br><br>If there is more than one entry for the questionnaires (GSRS and GI symptoms) on any given date or timepoint the worst score will be used for that date or timepoint as conservative approach.<br><br>If data were missing from one or more item in a GSRS questionnaire, the mean of the completed items in the same dimension will be used provided that more than half of the items in that dimension had been completed.<br><br>If data are missing from a PBC-40 domain (typically missed or duplicated answers) the whole domain should be discarded if <50% of items are completed. If >50% of responses are present, then the median value for the completed items in the domain will be ascribed to the missing item. |
| **Study Day** |
| Calculated as the number of days from start of initial study period date (Day 1) with no Day 0 defined:<br>Ref Date = Missing → Study Day = Missing<br>Ref Date < Date of Day 1 → Study Day = Ref Date – Date of Day 1<br>Ref Date ≥ Date of Day 1 → Study Day = Ref Date – date of Day 1 + 1 |

| Demographics |
|---|
| **Age** |
| GSK standard IDSL algorithms will be used for calculating age where date of birth (DOB) will be imputed as follows:<br>Any participant with a missing day will have this imputed as day '15'.<br>Any participant with a missing day and month will have this imputed as '30th June'.<br>Birth date will be presented in listings as 'YYYY'.<br>The Date of Birth will be assumed to be 30 June YYYY and age will be calculated at screening visit (Visit 1) using this assumed date of birth. |
| **Body Mass Index (BMI)** |
| Calculated as Weight (kg) / [Height (m)]$^2$ |

| Demographics |
|---|
| **Country** |
| A number of tables will be displayed by the subgroup Country (Japan only and The rest of the world). |
| General Displays note: |
| Use Planned Treatment (TRT01P or TRTP) for all outputs except for Safety and PK which will use Actual Treatment (TRT01A or TRTA). |

### 12.6.2. Study Population

| Treatment Compliance Number of Tablets |
|---|
| Compliance will be calculated based on the number of tablets dispensed and returned within the MSP (dispensed at V3, V4 and V5 and returned at V4, V5 and V6).<br><br>**MSP Compliance** =100 * (total number of tablets dispensed from V3 to V6 – total number of tablets returned from V3 to V6) / (expected number of tablets to be taken)<br>Where expected number of tablets to be taken = 6 tablets per day * (V6 –V3).<br><br>If a participant is lost to follow-up, compliance will be calculated up to last Visit attended, assuming no tablets were taken from any bottles dispensed at this visit. These participants should 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 MSP.<br><br>For example, if a participant does not return for V5 their compliance will be calculated as:<br><br>100 * (total number of tablets dispensed from V3 to V4 – total number of tablets returned from V3 to V4) / (expected number of tablets to be taken)<br>Where expected number of tablets to be taken = 6 tablets per day * (V4 –V3).<br><br>With a flag added to say that the participant dropped out after Visit 4.<br><br>Compliance based on number of tablets will also be calculated for the containers marked as AM (2 containers, 4 tablets a day) and PM bottles (1 container, 2 tablets a day) over the duration of the main study period (Exposure start date at V3 to Exposure end date at V6).<br><br>Compliance will be calculated similarly for the initial study period (V2 to V3) and final study period (V6 to V7).<br><br>Refer to Appendix 14 for further details. |
| **Compliance E-diary** |
| The compliance will be calculated for the period that participants are taking randomised study treatment during the Main Study period (MSP) from baseline (Visit 3) until the |

primary timepoint for the analysis (Visit 6).

For each day during this period, from the evening of Visit 3 until the morning of Visit 6, the following will be calculated:

CEQ1am = 1 if 'Have you taken your study medication for the morning' is Yes;
= 0 if No
CEQ2am = 1 if 'I will take my study treatment after I completed the diary' is chosen;
= 0 otherwise

CEQam = maximum(CEQ1am, CEQ2am) so = 1 if CEQ1am or CEQ2am = 1

Total CEQam = Number of days that CEQam = 1 between V3 + 1 and V6.

Total CEQpm will be calculated similarly (between V3 and V6-1) for the same questions in the PM diary entries.

**MSP AM Compliance** = 100 * (Total CEQam)/(V6-V3)
**MSP PM Compliance** = 100 * (Total CEQpm)/(V6-V3)
**MSP Compliance** =mean (MSP AM compliance, MSP PM Compliance)

**Extent of Exposure**

Number of days of exposure to study drug during MSP (V3 to V6) will be calculated based on the formula:
**Duration of Exposure in Days = Treatment Stop Date –Treatment Start Date + 1**
Participants who were randomised but did not report a treatment start date will be categorised as having zero days of exposure.

The cumulative dose will be based on the formula:
**Cumulative Dose = Sum of (Duration of Exposure x Total Daily Dose)**

If there are any treatment breaks during the study, exposure data will be adjusted accordingly. If the treatment break is greater than one day then the duration of exposure in days will be adjusted by excluding the treatment break.

If length of break>1 day then:
**Duration of Exposure in Days = Treatment Stop Date – Treatment Start Date + 1 – Duration of Breaks.**

### 12.6.3. Efficacy

**Mean Worst Daily Itch**

The primary analysis will be based on the Worst Itch NRS; this is completed twice per day (morning and evening) and the Worst Daily Itch Score is the maximum of the AM and PM responses.

| |
|---|
| Worst Daily Itch = Maximum (AM Worst Itch, PM Worst Itch). (possible scores 0-10)<br>Missing Data within day (where day is AM and PM on the same day):<br>Missing AM Worst Itch score then Worst Daily Itch = PM Worst Itch score<br>Missing PM Worst Itch score then Worst Daily Itch = AM Worst Itch score<br>Missing AM and PM Worst Itch score then Worst Daily Itch = Missing |
| Mean Worst Daily Itch at each visit is the mean of the Worst Daily Itch on the 7 days prior to the relevant visit<br>Mean Worst Daily Itch at:<br>Baseline (Visit 3) = average of worst daily itch V3 -1 to V3 -7.<br>Visit 4 = average of worst daily itch on V4 -1 to V4 -7<br>Visit 5 = average of worst daily itch on V5 -1 to V5 -7<br>Visit 6 = average of worst daily itch on V6 -1 to V6 -7<br>Visit 7 = average of worst daily itch on V7 -1 to V7 -7<br><br>If more than 3 of the 7 worst daily itch scores are missing in any one week then the mean worst daily itch will be set to missing.<br><br>If there is more than one entry in the e-diary data for 'Worst Itch' on any given day and timepoint (AM or PM) then the worst (the highest) score will be used from each duplicate record for that day and timepoint and it will be used in the derivation of mean worst daily itch. |
| Change from Baseline in Mean Worst Daily Itch from Baseline at Visit x<br>= Mean Worst Daily Itch at Visit x – Mean Worst Daily Itch at Visit 3 |
| **Centred Mean Worst Daily Itch at Baseline** |
| This will be used as a covariate in the analyses of Mean Worst Daily Itch Score. It will be calculated using only the participants with change from baseline in Mean Worst Daily Itch data available for analysis within the population of interest. |
| **Rolling Average Worst Daily Itch** |
| This is calculated for all participants in the ITT population. The analysis considers a rolling average of mean worst daily itch. This allows the daily fluctuation of the worst daily itch to be taken into account as well as ensuring that a clear picture of the mean worst daily itch can be seen throughout the study period.<br><br>For each patient the rolling average is calculated from day of V2+1 until day of V7.<br><br>For each day the rolling average of the worst daily itch from the 7 days prior to each day will be calculated. This will be calculated using only data available within the relevant study period.<br><br>**Initial Study Period (Day 2 to V3 inclusive):**<br>Day 2 = average worst daily itch from Day 1, |

Day3 = average of worst daily itch form Day 1 and Day 2
Day8 = average of worst daily itch from Day 1 to Day 7
Day 9 = average of worst daily itch from Day 2 to Day 8
…..
Day of V3 = average of worst daily itch from day of V3-1 to day of V3-7.

**Main Study Period (V3+1 to V6 inclusive):**
Day of V3 + 1 = average of worst daily itch from Day of V3
Day of V3 + 2 = average of worst daily itch from Day V3 to Day of V3 + 1
Day of V3 + 8 = average of worst daily itch from Day V3 to Day of V3 + 7
Day of V3 + 9 = average of worst daily itch from Day V3 + 1 to Day of V3 + 8
…….
Day of V6 = average of worst daily itch from day of V6-1 to day of V6-7.

**Final Study Period:**
Day of V6 + 1 = average of worst daily itch from Day of V6
Day of V6 + 2 = average of worst daily itch from Day V6 to Day of V6 + 1
Day of V6 + 8 = average of worst daily itch from Day V6 to Day of V6 + 7
Day of V6 + 9 = average of worst daily itch from Day V6 + 1 to Day of V6 + 8
…….
Day of V7 = average of worst daily itch from day of V7-1 to day of V7-7.

**Responder in Mean Worst Daily Itch**

Improvement in Mean Worst Daily Itch is a reduction in the score (10 is the highest and worst severity, 0 lowest and least severe).

Mean Worst Daily Itch Responders at Visit x:

Mean Worst Daily Itch Visit x Responder:
= 1, If Mean Worst Daily Itch Score at Vx is < 4
= 0, Otherwise

Mean Worst Daily Itch Percentage Improvement Responder
= 1, If ratio of baseline at Vx< 0.7 (i.e. at least a 30% reduction from baseline in worst daily itch score)
= 0, Otherwise

Mean Worst Daily Itch Improvement Responder
= 1, If change from V3 in mean worst daily itch at Vx ≤ - 2.
= 0, Otherwise

**Responder Days in Worst Daily Itch (Responder Days)**

Number of responder days between V3+1 and V6-1 where response is defined, according to the following 3 definitions, for each day as:

**Worst Daily Itch Responder:**
= 1, If Worst Daily Itch Score < 4

= 0, Otherwise
= missing if worst daily itch score is missing

**Worst Daily Itch Percentage Improvement Responder**
= 1, If ratio of baseline < 0.7 (i.e. at least a 30% reduction from baseline in worst daily itch score)
= 0, Otherwise
= missing if worst daily itch score is missing

**Worst Daily Itch Improvement Response**
= 1, If change from baseline in mean worst daily itch ≤ - 2.
= 0, Otherwise
= missing if worst daily itch score is missing

For each of these the baseline is taken as the Mean Worst Daily Itch at V3 (see above). The number of responder days is the total number of days where response is 1 between V3+1 and V6-1 inclusive.

The **percentage of responder** days during the MSP is then calculated, for each responder definition, as:

(Number of responder days / total number days from V3+1 to V6-1 for which a worst daily itch score was available) * 100

Note: days for which no worst daily itch score is available will not contribute to either the numerator or the denominator.

**Time to Worsening of Itch**

The improver population with improvement in Mean Worst Daily Itch at V6 i.e. those with:
Change in Mean Worst Daily Itch from Baseline ≤ -2 at Week 16 (i.e. Improved by at least 2 points between baseline and Week 16).

This is calculated in the subset of the ITT population with improved Mean Worst Daily Itch. The analysis considers a rolling average of mean worst daily itch. This allows the daily fluctuation of the worst daily itch to be taken into account as well as ensuring that the comparisons being made are equivalent. This is similar to the Rolling Average Worst Daily Itch form above, although there are different rules on what data are included in this calculation.

Data are restricted to use only the worst daily itch scores available in the Final Study Period (FSP) from V6+1 to V7-1.

In order for data to be comparable to mean worst daily itch score at Week 16 (the comparator used to determine if a participant has worsened), a 7 day moving average (MAV) of worst daily itch is calculated where at least 4 values must be non-missing.

MAVx = average of worst daily itch on Day x-3 to Day x+3 around the mid-point Day x from V6+1 to V7-1.

Note that for MAV at x=V6+1, there are no contributing values for the 3 days prior to Day x as these are not considered in the MSP. Similarly, for MAV at x=V7-1, there are no contributing values for the 3 days after Day x as these are not considered in the FSP.

Example:
Define FSP Day relative to V6 so that x days after V6 is FSP Day x.
So for example, MAV5 is calculated as the average worst daily itch scores over FSP Day 2 to 8.

For each MAVx, the median of the corresponding FSP Days is calculated (DMAVx).
So DMAX5 = median (2, 3, 4, 5, 6, 7, 8) = 5.

So, in the case of 7 contributing values and no missing data, DMAVx = x.

However, if for example, data on FSP Day 2 was missing then DMAX5 = median (3, 4, 5, 6, 7, 8) = 5.5.

Time of worsening is calculated as the first day (DMAVx) where the participants moving average around this point is higher than their Week 16 mean worst daily itch by 2 or more points:

MAVx - Mean Worst Daily Itch at V6 (Week 16) $\geq 2$

If the participant's MAV does not increase by 2 points (or more) by the end of the FSP then the time of worsening will be censored at the last DMAVx based on at least 4 evaluable days in FSP.

Example:
Consider a participant who does not worsen and attends for Visit 6 and Visit 7 on Days 82 and 112, respectively. The last MAV was calculated based on non-missing data on Days 108, 109 ,110, and 111 (which relative to V6 at Day 82 equates to FSP Day 26 to 29), the censored time of worsening value will be median (26, 27, 28, 29) = 27.5.

**PRO: Other eDiary Itch Questions**: **Mean Sleep and Fatigue Score**

Mean Sleep Score at each visit is the mean of the sleep scores (0-10 NRS) from the 7 days prior to the relevant visit. The Daily Sleep Score is obtained from the morning eDiary entries.

Mean Fatigue Score at each visit is the mean of the fatigue scores from the 7 days prior to the relevant visit. The Daily Fatigue Score is obtained from the evening eDiary entries.

Each of these are derived as for Mean Worst Daily Itch from Vx-1 to Vx-7 for Visit x.

If more than 3 of the 7 fatigue scores are missing in any one week then the Mean Fatigue Score is missing; a similar approach is taken for the Mean Sleep Score.

If there is more than one entry in the e-diary data for the sleep score on any given day then the worst score will be used for that day and it will be used in the derivation of mean

daily sleep score.

**PRO: Other eDiary Itch Questions: Mean Overall Itch**

Another question in the eDiary asks about 'overall or average itch' using a 0-10 NRS. This question is asked in both the morning and evening diaries and will be summarised by taking the average of the morning and evening scores each day and averaging this over the 7 days prior to each visit.

Average Overall Itch = Average (AM Overall itch, PM Overall itch). (possible scores 0-10)

Missing Data within day (where day is the AM and PM on the same date):

| | |
|---|---|
| Missing AM Overall Itch score then | Average Overall Itch = PM Overall Itch score |
| Missing PM Overall Itch score then | Average Overall Itch = AM Overall Itch score |
| Missing AM and PM Overall Itch score then | Average Overall Itch = Missing |

The Mean Overall Itch is then derived as for the Mean Worst Daily Itch, using the Average Overall Itch from Vx-1 to Vx-7 for Visit x.

**PRO: Other eDiary Itch Questions: Other eDiary Questions**

There are a number of other eDiary questions, listed below. They will be summarised over the 7 days prior to each visit in 2 ways; summarising the categorical answers and by deriving a numerical score.

**Average Response**

Table A below shows the scores that should be assigned to the answers for these questions.

Morning diary

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.

Evening Diary

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.

**Table A of Scoring for the Other eDiary questions**

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.

For each participant and question the average score over the 7 days prior to each visit will be calculated as per see Mean Sleep and Fatigue Score above.

As with all questions in the eDiary if the score for more than 3 of the 7 days is missing the Mean score will be missing.

**Most Frequent response**

For each participant/question

Most Frequent Response = Mode (response to question on Vx-1 to Vx-7)

If there is more than one mode the worst case will be taken as the response with the highest score (in the table above), with the exception of question A4 where the worst response is No (the response with the lowest score).

For example, if a participant answers to question A1 are as follows:

A little bit, Quite a bit, Quite a bit, A little bit, Not at all, Quite a bit, Quite a bit – the Most Frequent Response will be 'Quite a bit'

If a participant answers are:

A little bit, Quite a bit, Quite a bit, A little bit, Not at all, Quite a bit, A little bit – the Most Frequent Response will still be 'Quite a bit'. There will be 2 modes, 'A little bit' and 'Quite a bit', but 'Quite a bit' is the worst case for that question and so will be taken as the Most Frequent Response.

| **PTEA** |
|---|
| The responses to the PTEA (closed questions) will be displayed as categories and will not be converted to scores. |
| **Follow-Up Phone Call** |
| The follow-up phone call will ask about the participant's worst itch, worst fatigue and how much itching interferes with sleep on the 0-10 NRS. This will be presented as summary of the scores. |

The remaining questions have categorical responses which will be summarised using contingency tables.

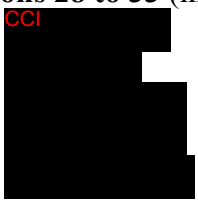
**PBC-40**

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 (summarised below).

**Question 1** is scored as:


**Question 2 to 27** (inclusive) are scored as:


**Questions 28 to 33** (inclusive) are scored as:




**Questions 34-39** (inclusive) are scored as:


**Question 40** is scored as:


For all Questions, an answer of CCI will be scored CCI.
For each visit the PBC-40 questionnaire should provide scores for 6 domains:
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 should be marked as missing. If 50% or

more of the corresponding question scores are complete the median of these responses should be imputed for the missing values within the domain before calculating the domain score.

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 33
PBC-Social = sum of scores of questions 29 31 32 and 34 to 40

If data are missing from a PBC-40 domain (typically missed or duplicated answers) the whole domain should be discarded if <50% of items are completed. If >50% of responses are present, then the median value for the completed items in the domain will be ascribed to the missing item.

**5-D Itch**

For each visit the 5-D itch score is broken down into 5 domains
Duration Score = Score for question on duration (on scale of 1 to 5)
Degree Score = Score for question on degree (on scale of 1 to 5)
Direction Score = Score for question direction (on scale of 1 to 5)
Disability Score = Maximum score given for all four questions under the Disability heading (sleep, leisure/social, housework/errands, work/school)
Distribution Score = Total number of areas of the body affected is obtained, which is then used to determine a Distribution score:
Total number of areas affected = 0-2 then Distribution Score = 1
Total number of areas affected = 3-5 then Distribution Score = 2
Total number of areas affected = 6-10 then Distribution Score = 3
Total number of areas affected = 11-13 then Distribution Score = 4
Total number of areas affected = 14-16 then Distribution Score = 5
Total 5-D itch score can then be taken as the total of the scores from each domain:
5-D Itch Score = Duration Score + Degree Score + Direction Score + Disability Score + Distribution Score.
If any domain score is missing then the Total 5-D Itch score will also be missing.

**EQ-5D-5L and EQ-VAS**

The EQ-5D is a standardised instrument used to measure health utility. It is a two-part self-assessment questionnaire, the first part being a descriptive system comprising five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/ depression, the second part requires the participant to indicate their perceived health on a visual analogue scale (EQ-VAS).

In the descriptive system for each of the five dimensions there are five levels of response ranging from ‘no problems’ to ‘extreme problems’ and reflecting the participant’s current health status. The responses are scored 1-5 as shown in Table B.

**Table B: EQ-5D scores**

| Dimension | Response | Score |
|---|---|---|
| Mobility | CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded. | |
| Self-care | | |
| Usual activities (e.g. work, study, housework, family or leisure activities) | | |
| Pain/discomfort | | |
| Anxiety/depression | | |

Ambiguous values (i.e. 2 boxes ticked for a single question) should be treated as missing.

A participant is classified into one of 3125 distinct health states (each being referred to by a 5-digit code based on the responses to each dimension in the descriptive system, e.g. 11111 or 55555).

This health state is then converted to a utility index score. EQ-5D-5L health state utility index scores summarize how good or bad health problems are on a scale anchored at CCI (CCI) and CCI (CCI). Health states considered worse than dead are given utility index scores <0.

The responses to each of the dimensions will be summarized at each assessment time for which there are adequate data.
The utility index score will be derived and summarized by pruritus severity and overall with summary statistics at each assessment time for which there are adequate data. Summary statistics on change from baseline will also be provided; overall and by baseline severity.

Due to concerns raised by independent quality assurance of the analytical methods used to create the 5L valuation set for England, NICE currently does not recommend using the 5L valuation set and that companies, academic groups and others preparing evidence submissions for NICE should use the 3L valuation set for reference-case analyses.

Utility values should be calculated by mapping the 5L descriptive system data onto the 3L valuation set. The nonparametric model mapping function developed by van Hout et

al. (2012) should be used. A more detailed description of the crosswalk model and methodology, the actual index values and an Excel based calculator can be downloaded from the EuroQol site: (https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/valuation-standard-value-sets/crosswalk-index-value-calculator/) the UK value set should be used. The Excel calculator for MS Windows is below (also available for Mac via the link above.

Other country specific value-sets are available (Denmark, France, Germany, Japan, the Netherlands, Spain, Thailand, UK, US and Zimbabwe) and future analyses may explore these.

**EQ-VAS**

The EQ-VAS requires the participant to indicate their perceived health on a visual analogue scale (EQ-VAS), where the endpoints are labelled 'CCI' anchored at CCI and 'CCI' anchored at CCI

The EQ-5D Thermometer will be summarized with summary statistics at each assessment time for which there are adequate data. Summary statistics on change from baseline will also be provided.

**PGI-S**

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

CCI

**PGI-S Response**

The response to the PGI-S question will also be summarised in terms of frequency and percentage at each Visit.

**PGI-C**

PGI-C: Score will be on a scale of 1-7 from the PGI-C question.

CCI

**PGI-C Response**

The response to the PGI-C question will also be summarised in terms of frequency and percentage at each Visit.

**BDI-II**

The Beck Depression Inventory –II consists of 21 questions, with possible responses being scored from [redacted] ([redacted]) to [redacted] ([redacted]). 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 |
|---|---|---|
| [redacted] | [redacted] | [redacted] |

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.

Question 18:

| Collection Code | Label | Item Score |
|---|---|---|
| [redacted] | [redacted] | [redacted] |

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.

**Actigraphy**

For each participant participating in the actigraphy sub-study, data will be provided for up to 5 records (nights) over 7 days in each of the 3 time-periods in which they participate. (Week 2 or Week 3, Week 9 or Week 10, Week 13 or Week 14).
If in a time period there are 3 or fewer records, then the data for that time period should be disregarded.
Otherwise all figures and analysis should use all of the data as provided.

For each of the 3 time periods for which data are available the average number of scratching events per hour will be calculated for the time period, for each participant, as:

Total number of scratching events for night j ($NS_j$)
= Average number of scratching events per hour on night j * duration of rest period on night j
$DR_j$ = Duration of rest period on night j.

Average number of scratching events per hour over the time period
= (sum of $NS_j$ in the time period) / (sum of $DR_j$ in the time period).

The average duration of scratching events over each time period will be calculated similarly.

**Risk Group**

The risk group strata is assigned for all participants at baseline visit (V3) for randomisation and is based on the lab results of the samples taken at visit 2 (ALP).

**High Risk Population**

The High Risk Population will be defined as those that were assigned to the High Risk strata for randomisation at V3. This population will be used to look at the change in disease progression markers in those at High Risk of disease progression.

**Restricted High Risk Population**

As part of a sensitivity analysis the High Risk population will be restricted to those confirmed to be still at High Risk of disease progression at V3, based on lab results and will additionally exclude anyone who was mis-stratified to the High Risk strata for randomisation.

Restricted High risk Population will be defined as:

| |
|---|
| Those assigned to High Risk strata for randomisation (V3) AND whose lab values at BOTH V2 and V3 meet the criteria below:<br><br>ALP at VX ≥ 1.67*ULN AND/OR total bilirubin at VX >ULN<br><br>The sensitivity analysis will only be undertaken if the Restricted High Risk population differs from the High Risk population. |
| **Disease Progression Risk Responders at each Visit** |
| For all of those in the High Risk Population.<br><br>= 1, If ALP at $V_X$ < 1.67 * ULN *AND* total bilirubin at $V_X$ ≤ ULN<br>= 0, Otherwise |

### 12.6.4. Safety

| |
|---|
| **GSRS Domain Score** |
| **GSRS Responses**<br>The responses to the individual GSRS questions asked at each visit will be displayed in terms of frequency and percentage at each Visit.<br><br>**GSRS Domain Score**<br>The responses to each question in the GSRS range from [redacted] ([redacted]) to [redacted] ([redacted]).<br><br>For each week, the GSRS questionnaire should provide scores for 5 domains:<br>Diarrhoea Syndrome = average of scores for questions 11, 12 and 14<br>Indigestion Syndrome = average of scores for questions 6, 7, 8 and 9<br>Constipation Syndrome = average of scores for questions 10, 13 and 15<br>Abdominal pain syndrome = average of scores for questions 1, 4 and 5<br>Reflux Syndrome = average of scores for questions 2 and 3<br>Total = average scores for questions 1-15<br><br>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 should be marked as missing. If 50% or more of the corresponding question scores are complete the mean of these responses should be imputed for the missing values within the domain before calculating the domain score.<br>For each Visit the GSRS scores will be the weekly questionnaire completed within 7 days prior to the visit.<br><br>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. |

**GI symptoms**

**Symptoms – Diarrhoea**

There are 3 diarrhoea questions within the weekly eDiary, with categorical 5-point responses.

Response will be provided to the weekly eDiary diarrhoea questions each week after start of ISP. The weekly eDiary diarrhoea responses for a visit will be those provided within 7 days prior to the visit.

If there are multiple entries for that same question for the questionnaires (GI symptoms) on the same date the worst score will be used for that date as conservative approach.

**Symptoms – Abdominal Pain**

There are 2 abdominal pain questions within the weekly eDiary, one with categorical 5-point response and one with 0-10 NRS.

Response will be provided to the weekly eDiary abdominal pain questions each week after start of ISP. The weekly eDiary abdominal pain responses for a visit will be those provided within 7 days prior to the visit.

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

**ECG Parameters: 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.

**ECG Parameters: 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:

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

**Adverse Events (AEs and AEs of Special Interest)**

Elevated ALT meeting stopping criteria, as per Section 8.1.1 of the protocol table 5.

| Post-Screening Lab Result | Actions |
|---|---|
| **Total bilirubin > 2x participant's screening AND >1.5x ULN** | **Stop study treatment immediately.**<br><br>**Discuss with Medical Monitor within 24 hours.**<br><br>Repeat liver chemistry testing (including AST, ALT, total bilirubin, ALP, and PT/INR) **within 24-72 hours** where possible **but within 1 week,** and thereafter **at least weekly** until values stabilize or fall below the criteria for increased monitoring.<br><br>Refer to **Protocol Appendix 4** for further follow-up and monitoring requirements. |
| **Total bilirubin >3x ULN** | |
| **ALT > 3x participant's screening ALT result AND > 5x ULN** | |
| **ALT >8x ULN of ALT** | |

In addition, study treatment will be stopped for a participant if there are liver chemistry elevations which, in the opinion of the investigator, are not attributable to the participant's underlying PBC, or if there is worsening liver chemistry associated with appearance of new symptoms which may typically be associated with drug-induced liver injury (including fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or eosinophilia [>5%]).

As the ALT would naturally fluctuate in these subjects, the AESIs are the ALTs that have met the stopping criteria above. If the stopping criteria has been met then the liver event form should be completed.

**Diarrhoea and Abdominal Pain**

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 Abdominal Pain AEs and diarrhoea AEs of special interest 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 provided in a separate document 'TOI Diarrhoea_MedDRA.xlsx'

**Other Adverse Events**

At blinded data review other adverse events may be identified for further investigation.

| 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.<br>Example 1: 2 Decimal Places = '< x ' becomes x – 0.001 e.g. <0.05 becomes 0.049<br>Example 2: 1 Decimal Place = '< x' becomes x – 0.01 e.g. <0.2 becomes 0.19<br>Example 3: 0 Decimal Places = '> x' becomes x + 0.1 e.g. >99 becomes 99.1 |

### 12.6.5. Pharmacokinetic

| Endpoints detailed in Ssection in 9.1.1.1 |
|---|

### 12.6.6. Pharmacodynamic Biomarker

| Endpoints and variables are detailed in Section 10.1.1 |
|---|

## 12.7. Appendix 7: Reporting Standards for Missing Data

### 12.7.1. Premature Withdrawals

| Element | Reporting Detail |
|---|---|
| General | • Participant study completion was defined as in the Protocol Section 5.3.<br>• Withdrawn participants will not be replaced in the study.<br>• All available data from participants who were withdrawn from the study will be listed and all available planned data will be included in summary tables and figures, unless otherwise specified.<br>• Withdrawal visits will be slotted as per Appendix 3: Assessment Windows or will be summarised as withdrawal visits. |

### 12.7.2. Handling of Missing Data

| Element | Reporting Detail |
|---|---|
| General | • Missing data occurs when any requested data is not provided, leading to blank fields on the collection instrument:<br>• These data will be indicated by the use of a "blank" in participant listing displays. Unless all data for a specific visit are missing in which case the data is excluded from the table.<br>• Answers such as "Not applicable" and "Not evaluable" are not considered to be missing data and should be displayed as such. |
| Outliers | • Any participants excluded from the summaries and/or statistical analyses will be documented along with the reason for exclusion in the clinical study report. Footnotes should be added to the outputs, where appropriate. |

#### 12.7.2.1. Handling of Missing and Partial Dates

| Element | Reporting Detail |
|---|---|
| General | Partial dates will be displayed as captured in participant listing displays. |
| Adverse Events | The eCRF allows for the possibility of partial dates (i.e., only month and year) to be recorded for AE start and end dates; that is, the day of the month may be missing. In such a case, the following conventions will be applied for calculating the time to onset and the duration of the event:<br>Missing Start Day: First of the month will be used unless this is before the start date of randomised study treatment; in this case the study treatment start date will be used and hence the event is considered On-treatment.<br>Missing Stop Day: Last day of the month will be used, unless this is after the stop date of randomised study treatment; in this case the study treatment stop date will be used.<br>Completely missing start or end dates will remain missing, with no imputation applied. Consequently, time to onset and duration of such events will be missing. |

| Element | Reporting Detail |
|---|---|
| Concomitant Medications/Medical History | Partial dates for any concomitant medications and medical history recorded in the CRF will be imputed using the following convention:<br>If the partial date is a start date, a '01' will be used for the day and 'Jan' will be used for the month<br>If the partial date is a stop date, a '28/29/30/31' will be used for the day (dependent on the month and year) and 'Dec' will be used for the month. Unless after the study completion the end date will used. For example, if a participant ended study on 21 May 2019 then stop date will end of study date.<br><br>The recorded partial date will be displayed in listings. |
| Handling of values outside limit of quantification | Data marked as below a value (e.g. < X) will be replaced with 0.5*X for any summaries and analysis of this data. A footnote should be included in any tables to indicate the imputation that has been made. |
| Handling of Missing Data for Statistical Analysis | Randomised participants that withdraw from treatment early will be followed-up at each Visit as per the protocol with the aim of reducing the level of dropouts. Only those who withdraw from the study entirely or are lost to follow-up will not complete the Study Visits.<br><br>There are four possible ways that data can be missing:<br>• Withdrawn/lost to follow-up<br>• Missed Visit<br>• Missing data within a Visit<br>• Missing data within an instrument/measure<br><br>The critical time points for the majority of primary and secondary endpoints are the baseline measurement (Visit 3) and the Week 16 measurement (Visit 6), and for eDiary data the 7 days prior to these visits.<br><br>It is expected that the level of missing data will be less than 10%. Prior to breaking the blind, the data will be reviewed, examining the pattern and percentage of missing data. If the level of missing data is greater than 10%, different methods of dealing with this missing data will be investigated, including:<br>• Use all available data to derive the endpoints<br>In general, Mean Worst Daily Itch is missing if the worst daily itch is missing for more than 3 days out of 7. If this option were to be chosen the Mean Worst Daily Itch would be using all data available within the relevant 7 days.<br>• Fit Mixed Model Repeated Measures (MMRM) using all data available at Visits 3, 4 and 5 to investigate the change from baseline.<br>• Use Multiple Imputation Techniques to replace the missing data |

| Element | Reporting Detail |
|---|---|
| | For some of the derived endpoints, if the level of missing raw data reaches a certain level then the endpoint is forced to be missing. In these instances it may be possible to make use of the unused raw data, depending on the method that is chosen to deal with the missing data.<br><br>Last Observation Carried Forward (LOCF) will not be used to replace missing data. |

## 12.8. Appendix 8: Values of Potential Clinical Importance

### 12.8.1. Laboratory Values

| Haematology | | | | |
|---|---|---|---|---|
| **Laboratory Parameter** | **Units** | **Category** | **Clinical Concern Range** | |
| | | | **Low Flag (< x)** | **High Flag (>x)** |
| Haematocrit | Ratio of 1 | Male | | 0.54 |
| | | Female | | 0.54 |
| | | Δ from BL | ↓0.075 | |
| Haemoglobin | g/L | Male | | 180 |
| | | Female | | 180 |
| | | Δ from BL | ↓25 | |
| Lymphocytes | $x10^9$/ L | | 0.8 | |
| Neutrophil Count | $x10^9$/ L | | 1.5 | |
| Platelet Count | $x10^9$/ L | | 100 | 550 |
| While Blood Cell Count (WBC) | $x10^9$/ L | | 3 | 20 |
| Red Blood Cell Count (RBC) | | | NA | NA |

| Clinical Chemistry | | | | |
|---|---|---|---|---|
| **Laboratory Parameter** | **Units** | **Category** | **Clinical Concern Range** | |
| | | | **Low Flag (< x)** | **High Flag (>x)** |
| Albumin | g/L | | 30 | |
| Albumin | | CFB | 20% decrease | |
| Calcium | mmol/L | | 2 | 2.75 |
| Creatinine | umol/L | Δ from BL | | ↑ 26.5 |
| Glucose (fasting) | mmol/L | | 3 | 9 |
| Magnesium | mmol/L | | 0.5 | 1.23 |
| Potassium | mmol/L | | 3 | 5.5 |
| Sodium | mmol/L | | 130 | 150 |
| eGFR (CKD-EPI) | mL/min/ $1.73m^2$ | | <45 | - |
| BUN | | | NA | NA |

| Liver Function | | | |
|---|---|---|---|
| **Test Analyte** | **Units** | **Category** | **Clinical Concern Range** |
| ALT | U/L | % Δ from BL | ↑ 50% |
| AST | U/L | % Δ from BL | ↑ 50% |
| ALP | U/L | % Δ from BL | ↑ 50% |
| T Bilirubin | µmol/L | % Δ from BL | ↑ 50% |
| T. Bilirubin + ALT | µmol/L | High | 2xULN T. Bilirubin |

| Liver Function | | | |
|---|---|---|---|
| **Test Analyte** | **Units** | **Category** | **Clinical Concern Range** |
| | U/L | | + ≥ 5x ULN ALT |
| GGT | U/L | | NA |

### 12.8.2. ECG

| ECG Parameter | Units | Clinical Concern Range | |
|---|---|---|---|
| | | Lower | Upper |
| **Absolute** | | | |
| Absolute QTc Interval | msec | | > 450 [1] |
| Absolute PR Interval | msec | < 110 [1] | > 220 [1] |
| Absolute QRS Interval | msec | < 75 [1] | > 110 [1] |
| **Change from Baseline** | | | |
| Increase from Baseline QTc | msec | > 60 [1] | |

### 12.8.3. Vital Signs

| Vital Sign Parameter (Absolute) | Units | Clinical Concern Range | |
|---|---|---|---|
| | | Lower | Upper |
| Systolic Blood Pressure | mmHg | < 85 | > 160 |
| Diastolic Blood Pressure | mmHg | < 45 | > 100 |
| Heart Rate | bpm | < 40 | > 110 |

## 12.9. Appendix 9: Population Pharmacokinetic (PopPK) Analyses

### 12.9.1. Population Pharmacokinetic (PopPK) Dataset Specification

[NA]

### 12.9.2. Population Pharmacokinetic (PopPK) Methodology

[NA]

## 12.10. Appendix 10: Pharmacokinetic / Pharmacodynamic Analyses

### 12.10.1. Pharmacokinetic / Pharmacodynamic Dataset Specification

[NA]

## 12.11. Appendix 11: Abbreviations & Trade Marks

### 12.11.1. Abbreviations

| Abbreviation | Description |
|---|---|
| ADaM | Analysis Data Model |
| ADCM | Data Structure for Adverse Event Analysis |
| ADAE | Data Structure for Adverse Event Analysis |
| AE | Adverse Event |
| AIC | Akaike's Information Criterion |
| ALP | Alkaline Phosphatase |
| ALT | Alanine Aminotransferase |
| ANCOVA | Analysis of Covariance |
| AR(1) | Autoregressive |
| AST | Aspartate Aminotransferase |
| BDI-II | Beck Depression Inventory-II |
| BID | Twice a Day |
| BL | Baseline |
| BLC | Baseline Change |
| BP | Blood Pressure |
| bpm | beats per minute |
| BUN | Blood Urea Nitrogen |
| CDISC | Clinical Data Interchange Standards Consortium |
| CI | Confidence Interval |
| CS | Compound Symmetry |
| C4 | 7-α-hydroxy-4-cholesten-3-one |
| DBF | Database Freeze |
| Cmax | Maximum concentration |
| Cave | Average concentration |
| DOB | Date of Birth |
| DP | Decimal Places |
| ECG | Electrocardiogram |
| eCRF | Annotated Case Record Form |
| eDiary | Electronic Diary |
| ELF | Enhanced Liver Fibrosis |
| eGFR | Estimated glomerular filtration rate using (CKD-EPI equation) |
| EQ-5D-5L | EuroQOL 5D-5L |
| FGF-19 | Fibroblast Growth Factor -19 |
| FOBT | Fecal Occult Blood Test |
| FSP | Final Study Period |
| FU | Follow-Up |
| GGT | Gamma-Glutamyl Transpeptidase |
| GI | Gastrointestinal |
| GSRS | Gastrointestinal Symptoms Rating Scale |
| H0 | Primary Null Hypothesis |
| H1 | Alternate Hypothesis |
| HDL | High Density Lipoprotein |
| HR | Heart Rate |
| IA | Interim Analysis |
| IAP | Interim Analysis Plan |

| Abbreviation | Description |
|---|---|
| IBAT | Ileal Bile Acid Transporter |
| ICF | Informed Consent Form |
| ICH | International Conference on Harmonisation |
| IDSL | Integrated Data Standards Library |
| IP | Investigational Product |
| ISP | Initial Study Period |
| ITT | Intent-To-Treat |
| GSK | GlaxoSmithKline |
| GUI | Guidance |
| L | Litres |
| LDL | Low-Density Lipoprotein |
| LOCF | Last Observation Carries Forward |
| LS | Least Squares |
| MAV | Moving Average |
| Max | Maximum |
| MCV | Mean Corpuscular Volume |
| MCH | Mean Corpuscular Hemoglobin |
| mg | milligram |
| Min | Minimum |
| MMRM | Mixed Model Repeated Measures |
| msec | milliseconds |
| MSP | Main Study Period |
| NRS | Numerical Rating Scale |
| PBC | Primary Biliary Cholangitis |
| PCI | Potential Clinical Importance |
| PD | Pharmacodynamic |
| PDMP | Protocol Deviation Management Plan |
| PK | Pharmacokinetic |
| PGI-C | Patient Global Impression of Change |
| PGI-S | Patient Global Impression of Severity |
| PP | Per Protocol |
| PRO | Patient Reported Outcome |
| PT | Preferred Term |
| PT/INR | Prothrombin Time/International Normalised Ratio |
| QC | Quality Control |
| QD | Once a Day |
| QTcF | Frederica's QT Interval Corrected for Heart Rate |
| QTcB | Bazett's QT Interval Corrected for Heart Rate |
| RAP | Reporting & Analysis Plan |
| RAMOS | Randomisation & Medication Ordering System |
| RBC | Red Blood Cells |
| RTF | Rich Text File |
| SAC | Statistical Analysis Complete |
| SDTM | Study Data Tabulation Model |
| SD | Standard Deviation |
| SI | Standard |
| SOA | Schedule of Activities |
| SOC | System Organ Class |
| TFL | Tables, Figures & Listings |

| Abbreviation | Description |
|---|---|
| U | Universal Unit |
| UDCA | Ursodeoxycholic acid |
| UK | United Kingdom |
| ULN | Upper Limit of Normal |
| USA | United States of America |
| VAS | Visual Analogue Scale |
| WASO | Wake After Sleep Onset |
| WBC | White Blood Cells |
| µg | Micrograms |
| µmol | Micromoles |
| V | Visit |

### 12.11.2. Trademarks

| Trademarks of the GlaxoSmithKline Group of Companies |
|---|
| None |

| Trademarks not owned by the GlaxoSmithKline Group of Companies |
|---|
| [Phoenix WinNonlin] |
| [SAS] |

# 12.12. Appendix 12: List of Data Displays

## 12.12.1. Data Display Numbering

The following numbering will be applied for RAP generated displays:

| Section | Tables | Figures |
|---|---|---|
| Study Population | 1.1 to 1.39 | |
| Efficacy | 2.1 to 2.110 | 2.1 to 2.53 |
| Safety | 3.1 to 3.87 | 3.1 to 3.15 |
| Pharmacokinetic | 4.1 to 4.12 | 4.1 to 4.18 |
| Pharmacodynamic and / or Biomarker | 5.1 to 5.30 | 5.1 to 5.3 |
| **Section** | **Listings** | |
| ICH Listings | 1 to 31 | |
| Other Listings | 32 to 61 | |

## 12.12.2. Mock Example Shell Referencing

Non IDSL specifications will be referenced as indicated and if required example mock-up displays provided in Appendix 13: Example Mock Shells for Data Displays.

| Section | Figure | Table | Listing |
|---|---|---|---|
| Study Population | POP_Fn | POP_Tn | POP_Ln |
| Efficacy | EFF_Fn | EFF_Tn | EFF_Ln |
| Safety | SAFE_Fn | SAFE_Tn | SAFE_Ln |
| Pharmacokinetic | PK_Fn | PK_Tn | PK_Ln |
| Pharmacodynamic and / or Biomarker | PD_Fn | PD_Tn | PD_Ln |

**NOTES:**

- Non-Standard displays are indicated in the 'IDSL / Example Shell' or 'Programming Notes' column as '[Non-Standard] + Reference.'

## 12.12.3. Deliverables

| Delivery [Priority] [1] | Description |
|---|---|
| HLR | Headline Results |
| SAC | Final Statistical Analysis Complete |
| Country | By Japan only and by The rest of the world |

**NOTES:**

1. Indicates priority (i.e. order) in which displays will be generated for the reporting effort

## 12.12.4. Study Population Tables

| Study Population Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 1.1. | Randomised | ES1 | Summary of Participant Disposition | As per shell | SAC |
| 1.2. | Initial | ES4 | Summary of Participant Disposition at Each Study Phase | Where study phases are:<br>Screening (V1, prior to V2).<br>ISP If the Participant is assigned single-blind placebo at V2 they have started the ISP.<br>MSP: If the Participant is assigned to the double-blind treatment at V3 they have entered the MSP.<br>FSP: If the Participant is assigned the placebo at V6 they have entered the FSP.<br>Follow-up Participant completes V7 they have entered follow-up (which ends at time of V8 – follow-up phone call).<br>Footnote:<br>The Initial Study period starts on Visit 2 and ends on Visit 3.<br>The Main Study Period starts on Visit 3 and ends on Visit 6.<br>The Final Study Period starts on Visit 6 and ends on Visit 7.<br>Follow-up runs to V8. | HLR |
| 1.3. | Enrolled | ES6 | Screening Status and Summary of Reasons for Screen Failure | To produce a table describing, for all those that were enrolled into the study, the reasons for screen failure<br>Add a footnote to table: Participants are classed as screening failures if they do not pass into the ISP. | SAC |

| Study Population Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 1.4. | Initial | ES6 | Summary of Reasons for Baseline Failure | To produce a table describing, for all those that were entered into the ISP, the reasons for randomisation failure.<br>As per ES6 but with:<br>'Randomisation Status' rather than 'Screening Status'<br>'Entered Initial Study Period' rather than 'Enrolled'<br><br>Add a footnote to table: Participants are classed as randomisation failures if they do not pass into the MSP. | SAC |
| 1.5. | Safety | SD1 | Treatment Status and Summary of Reasons for Discontinuation of Study Treatment During the Main Study Period | | SAC |
| 1.6. | Initial | ES10 | Treatment Status and Summary of Reasons for Discontinuation of Placebo Run-in During the Initial Study Period | This will be shown separately for the initial study period and will show anyone who has stopped taking placebo during either of these periods | SAC |
| 1.7. | Safety | SD1 | Treatment Status and Summary of Reasons for Discontinuation of Placebo Treatments During the Final Study Period | This will be shown separately for the final study period and will show anyone who has stopped taking placebo during either of these periods<br>Programming notes: only the placebo during the final study period column should be showed for this table | SAC |
| 1.8. | Initial | SD1 | Summary of Reasons for Withdrawal from the Study | Please add footnote: Note: During the entire study period<br><br>Programming note: during the entire study period | SAC |
| 1.9. | Enrolled | NS1 | Summary of Number of Participants by Centre | As per shell NS1 with the addition of column 'not entered into ISP' as well as 'not randomised', where not entered into ISP is a screening failure and not randomised is a randomisation failure | SAC |

| Study Population Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 1.10. | Randomised | DV1 | Summary of Important Protocol Deviations | | SAC |
| 1.11. | Enrolled | Pop_T1 | Summary of Populations | Summary of N in each of the populations.<br>Calculate the percentages for the Restricted High Risk population based on the ITT population.<br>Programming notes: remove the actigraphy pop from the table and add the PK pop.<br>Please see Shell for update. | HLR |
| 1.12. | ITT | Pop_T4 | Summary of Risk Group Miss-stratification | | SAC |
| 1.13. | ITT | Pop_T5 | Summary of Changes in Risk Group Prior to Main Study Period | | SAC |
| 1.14. | Safety | DM1 | Summary of Demographic Characteristics | | HLR |
| 1.15. | Safety | DM11 | Summary of Age Ranges | | SAC |
| 1.16. | Safety | DM5 | Summary of Race and Racial Combinations | | SAC |
| 1.17. | Safety | DM6 | Summary of Race and Racial Combinations Details | | SAC |
| 1.18. | Safety | SU1 | Summary of Substance Use | | SAC |
| 1.19. | Safety | CM1 | Summary of Concomitant Medication ATC Categories by Ingredient | | SAC |
| 1.20. | Safety | CM1 | Summary of Previous Treatment for PBC | | SAC |
| 1.21. | Safety | MH1 | Summary of Past Medical Conditions | | SAC |
| 1.22. | Safety | MH1 | Summary of Current Medical Conditions | | SAC |
| 1.23. | Safety | Pop_T6 | Summary of PBC History | All information on MEDHX_PBC eCRF page, including Duration of PBC, Duration of Pruritus, PBC Diagnostic Criteria and Liver Transplant History.<br>To calculate duration please use the date of onset of symptoms/diagnosis up to Screening, impute missing dates per the RAP rules for Medical History. Liver Transplant History will be reported as Yes/No. | SAC |

| Study Population Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 1.24. | Safety | LB1 | Summary of Liver Chemistry at Screening | It would be a LB summary which you would subset to the liver chemistry tests. | SAC |
| 1.25. | Safety | EX1 | Summary of Exposure | For the Main Study Period – time on active treatment – from V3 to V6 (inclusive) | SAC |
| 1.26. | Safety | Pop_T3 | Summary of Treatment Compliance (defined by e-diary) in Main Study Period (AM, PM and Total) | For the Main Study Period – compliance in terms of e-diary for morning (AM), evening (PM) and Total – from V3 to V6 (inclusive). | SAC |
| 1.27. | Safety | Pop_T3 | Summary of Treatment Compliance (defined by tablets taken) in Main Study Period (AM, PM and Total) | For the Main Study Period – compliance in terms of tablets taken for morning (AM), evening (PM) and Total – from V3 to V6 (inclusive)<br><br>Footnote should read:<br>Note: compliance calculations are based on the number of tablets dispensed and returned. | SAC |
| 1.28. | ITT | Pop_T2 | Summary of Participants Attendance at Clinic Visits and Participants with Evaluable Diary Itch Data Evaluable Mean Worst Daily Itch | | SAC |
| 1.29. | ITT | COV1 | Summary of Covariates for Analysis of Mean Worst Itch Score | Summarise the following covariates, strata, and subgroups.<br>Footnote "Risk of Disease Progression was used as assigned for randomisation<br>See Section 5.4 for the full list | HLR |
| **Country** | | | | | |
| 1.30. | Randomised | ES1 | Summary of Participant Disposition by Country | As per shell | SAC |
| 1.31. | ITT | DV1 | Summary of Protocol Deviations by Country | | SAC |
| 1.32. | Enrolled | Pop_T1 | Summary of Populations by Country | Summary of N in each of the populations. | SAC |
| 1.33. | Safety | DM1 | Summary of Demographic Characteristics by Country | | SAC |

| Study Population Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 1.34. | Safety | Pop_T6 | Summary of PBC History by Country | All information on MEDHX_PBC eCRF page, including Duration of PBC, Duration of Pruritus, PBC Diagnostic Criteria and Liver Transplant History.<br>To calculate duration please use the date of onset of symptoms/diagnosis up to Screening, impute missing dates per the RAP rules for Medical History. Liver Transplant History will be reported as Yes/No. | SAC |
| 1.35. | Safety | LB1 | Summary of Liver Chemistry at Screening by Country | | SAC |
| 1.36. | Safety | EX1 | Summary of Exposure by Treatment Group by Country | For the Main Study Period – time on active treatment – from V3 to V6 (inclusive) | SAC |
| 1.37. | Safety | Pop_T3 | Summary of Treatment Compliance (defined by e-diary) in Main Study Period (AM, PM and Total) by Country | For the Main Study Period – compliance in terms of e-diary for morning (AM), evening (PM) and Total – from V3 to V6 (inclusive). | SAC |
| 1.38. | Safety | Pop_T3 | Summary of Treatment Compliance (defined by tablets taken) in Main Study Period (AM, PM and Total) by Country | For the Main Study Period – compliance in terms of tablets taken for morning (AM), evening (PM) and Total – from V3 to V6 (inclusive)<br><br>Footnote should read:<br>Note: compliance calculations are based on the number of tablets dispensed and returned. | SAC |
| 1.39. | ITT | COV1 | Summary of Covariates for Analysis of Mean Worst Itch Score by Country | Summarise the following covariates, strata, and subgroups.<br>Footnote “Risk of Disease Progression was used as assigned for randomisation<br>See Section 5.4 for the full list | SAC |

### 12.12.5. Efficacy Tables

| Efficacy: Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Mean Worst Daily Itch Score** | | | | | |
| 2.1. | ITT | Eff_T1 | Summary of Mean Worst Daily Itch Score. | For V3 to V7 (inclusive)<br>Note: N, n, Mean, SD, Min, Max, Median by Treatment Group and Visit | SAC |
| 2.2. | ITT | Eff_T1 | Summary of Change from Baseline in Mean Worst Daily Itch Score | For V3 to V7 (inclusive) | HLR |
| 2.3. | ITT | Eff_T1 | Summary of Percent Change from Baseline in Mean Worst Daily Itch Score | For V3 to V7 (inclusive) | SAC |
| 2.4. | ITT | Eff_T2 | Preliminary ANCOVA Analysis of Change from Baseline in Mean Worst Daily Itch Score at Week 16 (V6). | Mean difference for each treatment group from placebo, along with mean difference between 90 mg BID and 180 mg QD.<br>See Appendix 14 for updated shell Eff_T2 | HLR |
| 2.5. | ITT | Eff_T3 | Preliminary Linear Model of Change from Baseline in Mean Worst Daily Itch Score at Week 16 (V6). | Mean difference for each total daily dose level from placebo. Note inclusion of the 90 mg BID group (at total daily dose of 180 mg) depends on 2.4. | SAC |
| 2.6. | ITT | Eff_T3a | See Appendix 14 | | SAC |
| 2.7. | ITT | Eff_T4a (for ANCOVA) or Eff_T4b | Primary Analysis of Change from Baseline in Mean Worst Daily Itch Score at Week 16 (V6). | This will only be displayed for the final chosen model (e.g. Emax 4-parameter, Emax 3-parameter, ANCOVA, Linear Model, Power Model).<br>If the ANCOVA model is the best model then this will be done for each treatment group in the study, otherwise it will be for total daily doses in the study.<br><br>Programming Note: Please see updated shell Eff_T4c under Appendix 14. | HLR |
| 2.8. | ITT | EFF_T18 | T-test analysis Change from Baseline in Mean Worst Daily Itch Score at Week 16 (V6) | T-test analysis comparison between placebo and each dose level | SAC |
| 2.9. | PP | Eff_T1 | Summary of Mean Worst Daily Itch Score. | For V3 to V7 (inclusive) | SAC |
| 2.10. | PP | Eff_T1 | Summary of Change from Baseline in Mean Worst Daily Itch Score. | For V3 to V7 (inclusive) | SAC |

| | | | | | |
|---|---|---|---|---|---|
| 2.11. | PP | Eff_T4a (ANCOVA) or Eff_T4b | Analysis of Change from Baseline in Mean Worst Daily Itch Score at Week 16 (V6). | The final model chosen as part of the model selection process for the ITT population will be fitted using the PP population.<br><br>This will only be displayed for the final chosen model (e.g. Emax 4-parameter, Emax 3-parameter, ANCOVA, Linear Model, Power Model).<br>If the ANCOVA model is the best model then this will be done for each treatment group in the study. | SAC |
| 2.12. | ITT | Eff_T7 | MMRM Analysis of Change from Baseline in Mean Worst Daily Itch Score for each Treatment Group. | This will be done using data Visits 4 -7. This will be done for each treatment group.<br>Mean difference for each treatment group from placebo at each visit as well as mean difference between 90 mg BID and 180 mg QD treatment groups. | SAC |
| 2.13. | ITT | Eff_T8 | Analysis of Change from Baseline in Mean Worst Daily Itch Score at Week 16 (V6). ANCOVA Investigation of Risk Group. | The model fitted here will include Risk Group, Treatment Group and Risk Group-Treatment Group interaction | SAC |
| 2.14. | ITT | Eff_T14 | Analysis of Change from Baseline in Mean Worst Daily Itch Score at Week 16 (V6). Estimated Bias Associated with Ignoring Impact of Risk Group | ONLY if Risk Group is significant in ANCOVA | SAC |
| 2.15. | ITT | Eff_T15 | Analysis of Change from Baseline in Mean Worst Daily Itch Score at Week 16 (V6). ANCOVA Results: Impact of covariates | Results for ANCOVA models fitting each covariate in Section 5.4 one at a time will be displayed. | SAC |
| **Itch Responders** | | | | | |
| 2.16. | ITT | Eff_T9 | Summary of Mean Worst Daily Itch Responders | Frequency and Percentage of responders.<br><br>The three definitions should be shown in the one table.<br><br>For V4 to V7 (inclusive) | HLR |
| 2.17. | ITT | Eff_T10 | Logistic Regression Analysis of Mean Worst Daily Itch Responders at Week 16 (V6) | Odds of response at Week 16, and odds ratio of reaching responder criteria for each treatment group vs. placebo, for each placebo definition. | SAC |

| | | | | | |
|---|---|---|---|---|---|
| **Responder Days** | | | | | |
| 2.18. | ITT | Eff_T1 | Summary of Responder Days in Worst Daily Itch During the Main Study Period | By treatment group.<br>Include all 3 responder definitions in one table with column added to left of table to identify the response definition<br>***1) Worst Daily Itch < 4***<br>***2) Improvement in Mean Worst Daily Itch from Baseline ≥ 2***<br>***3) Improvement in Worst Daily Itch from Baseline ≥30%***<br>Note: Percentage and number of Days of Responder Days in Mean Worst Daily Itch: number of days response V3+1 and V6-1/ number of days mean worst daily itch scores available | SAC |
| 2.19. | ITT | Eff_T4a | ANCOVA Analysis of Responder Days in Mean Worst Daily Itch at Week 16 (V6) | Include all 3 responder definitions in one table with column added to left of table to identify the response definition<br>***1) Mean Worst Daily Itch < 4***<br>***2) Improvement in Mean Worst Daily Itch from Baseline ≥ 2***<br>***3) Improvement in Mean Worst Daily Itch from Baseline ≥30%***<br>Note: Percentage of Responder Days in Mean Worst Daily Itch: number of days response V3+1 and V6-1/ number of days mean worst daily itch scores available<br><br>Programming notes: please remove number of days parameter and keep %of days only for this table. | SAC |

| | | | | | |
|---|---|---|---|---|---|
| **Time to Worsening of Mean Worst Daily Itch** | | | | | |
| 2.20. | Responders (>=2 point improvement) | Eff_T1 | Summary of Time to Worsening of Mean Worst Daily Itch following Visit 6 in those that have improved between Baseline and Visit 6. | Footnotes as below:<br>Note: Improvement is defined as those that have a change from baseline in mean worst daily itch of at least -2.<br>Note: A participant is defined as worsened at the time their rolling average Mean Worst Daily Itch is at least 2 higher than their Mean Worst Daily Itch at Visit 6.<br><br>Note: Treatment groups are only included in this analysis if they consist of at least 8 participants. | SAC |
| 2.21. | Responders (>=2 point improvement) | TTE7 | Time to Event Analysis for Time to Worsening of Mean Worst Daily Itch following Visit 6. | Instructions on how to derive the time to worsening is provided in the RAP. | SAC |
| **Other eDiary itch questions** | | | | | |
| 2.22. | ITT | Eff_T1 | Summary of Mean Daily Sleep Score | Footnote:<br>Note: Mean Daily Sleep Score is defined as the average of the daily sleep scores provided in the 7 days prior to the relevant visit. | SAC |
| 2.23. | ITT | Eff_T1 | Summary of Change from Baseline in Mean Daily Sleep Score | Note: Mean Daily Sleep Score is defined as the average of the daily sleep scores provided in the 7 days prior to the relevant visit. | SAC |
| 2.24. | ITT | Eff_T1 | Summary of Mean Daily Fatigue Score | Note: Mean Daily Fatigue Score is defined as the average of the daily fatigue scores provided in the 7 days prior to the relevant visit. | SAC |
| 2.25. | ITT | Eff_T1 | Summary of Change from Baseline in Mean Daily Fatigue Score | Note: Mean Daily Fatigue Score is defined as the average of the daily fatigue scores provided in the 7 days prior to the relevant visit. | SAC |
| 2.26. | ITT | Eff_T1 | Summary of Mean Overall Itch Score. | Footnote: Note: Mean Overall Itching Score is defined as the average of the mean of the morning and evening overall itching scores provided in the 7 days prior to the relevant visit. | SAC |
| 2.27. | ITT | Eff_T1 | Summary of Change from Baseline in Mean Overall Itch Score | Note: Mean Overall Itch Score is defined as the average of the mean of the morning and evening overall itch scores provided in the 7 days prior to the relevant visit. | SAC |
| 2.28. | ITT | Eff_T4a | ANCOVA Analysis of Change from Baseline in Mean Overall Itch, Mean Sleep and Mean Fatigue Scores at Week 16 (V6). | This will be displayed for both the mean daily sleep and mean daily fatigue and mean overall itching (on 0-10 NRS) scores in one table, paginated by endpoint. | SAC |
| 2.29. | ITT | Eff_T1 | Summary of Average Response to Categorical Itch Questions in eDiary (AM) | | SAC |

| | | | | | |
|---|---|---|---|---|---|
| 2.30. | ITT | Eff_T1 | Summary of Change from Baseline in Average Response to Categorical Itch Questions in eDiary (AM) | Between Baseline (V3) and Week 16 (V6) | SAC |
| 2.31. | ITT | Eff_T13 | Summary of Most Frequent Response for Categorical Itch Questions in eDiary (AM) | Repeat, paginated by question with full question in sub-title, for all questions in e-diary (AM) that are not on 0-10 NRS.<br>Footnote: Most Frequent Response is taken as the mode of responses for each participant in the 7 days prior to each Visit.<br>Footnote: If there is more than one mode the Most Frequent Response is taken to be the worst of the modes. | SAC |
| 2.32. | ITT | Eff_T17 | Shift Table of Most Frequent Response to Categorical Itch Questions in eDiary (AM) | Between Baseline (V3) and Week 16 (V6)<br>Footnote: Most Frequent Response is taken as the mode of responses for each participant in the 7 days prior to each Visit.<br>Footnote: If there is more than one mode the Most Frequent Response is taken to be the worst of the modes. | SAC |
| 2.33. | ITT | Eff_T1 | Summary of Average Response to Categorical Itch Questions in eDiary (PM) | Repeat, paginated by question with full question in sub-title, for all questions in e-diary (PM) that are not on 0-10 NRS. | SAC |
| 2.34. | ITT | Eff_T1 | Summary of Change from Baseline in Average Response to Categorical Itch Questions in eDiary (PM) | Between Baseline (V3) and Week 16 (V6) | SAC |
| 2.35. | ITT | Eff_T13 | Summary of Most Frequent Response to Categorical Itch Questions in eDiary (PM) | Repeat, paginated by question with full question in sub-title, for all questions in e-diary (PM) that are not on 0-10 NRS.<br>Footnote: Most Frequent Response is taken as the mode of responses for each participant in the 7 days prior to each Visit.<br>Footnote: If there is more than one mode the Most Frequent Response is taken to be the worst of the modes. | SAC |
| 2.36. | ITT | Eff_T17 | Shift Table of Most Frequent Response to Categorical Itch Questions in eDiary (PM) | Between Baseline (V3) and Week 16 (V6)<br>Footnote: Most Frequent Response is taken as the mode of responses for each participant in the 7 days prior to each Visit.<br>Footnote: If there is more than one mode the Most Frequent Response is taken to be the worst of the modes. | SAC |

| | | | | | |
|---|---|---|---|---|---|
| 2.37. | ITT | Eff_T1 | Summary of Itch, Fatigue and Sleep at Follow-up (Visit 8) | Table should include a summary of the scores give to worst itching, itching interfere with sleep and worst fatigue at the follow-up telephone call. | SAC |
| 2.38. | ITT | Eff_T13 | Summary of Response to Categorical Itch Questions at Follow-up (Visit 8) | Table should include counts and percentages of the responses to the categorical questions asked during the follow-up phone call | SAC |
| Validated Questionnaires | | | | | |
| 5-D Itch | | | | | |
| 2.39. | ITT | Eff_T1 | Summary of 5-D Itch Scores. | By treatment group and visit<br>Note: N, n, Mean, SD, Min, Max, Median<br>Paginated by domain, with domain (as below) in subtitle. Only for V2, V3, V6 and V7.<br>Duration, Degree, Direction, Disability, Distribution, Total 5-D Itch Score | SAC |
| 2.40. | ITT | Eff_T1 | Summary of Change from Baseline in 5-D Itch Scores. | Paginated by domain, with domain (as below) in subtitle. Only for, V6 and V7.<br>Duration, Degree, Direction, Disability, Distribution, Total 5-D Itch Score | SAC |
| 2.41. | ITT | Eff_T4a | ANCOVA Analysis of Change from Baseline in 5-D Itch at Week 16 (V6). | This will be displayed for each of the 5-D Itch domains as well as the overall score (with 95% CI). | SAC |
| PBC-40 | | | | | |
| 2.42. | ITT | Eff_T1 | Summary of PBC-40 Scores | Paginated by domain, with domain (as below) in subtitle.<br>Itch, Fatigue, Symptoms, Cognitive, Social, Emotional<br>For V1 to V7 (inclusive)<br><br>Note: Summary (N, n, Mean, SD, Min, Max, Median) by Treatment Group | SAC |
| 2.43. | ITT | Eff_T1 | Summary of Change from Baseline in PBC-40 Scores | Paginated by domain, with domain (as below) in subtitle.<br>Itch, Fatigue, Symptoms, Cognitive, Social, Emotional<br>For V4 to V7 (inclusive)<br>Note: Summary (N, n, Mean, SD, Min, Max, Median) by Treatment Group | SAC |
| 2.44. | ITT | Eff_T4a | ANCOVA Analysis of Change from Baseline in PBC-40 at Week 16 (V6). | This will be displayed for each of the 6 PBC-40 domains. | SAC |
| Other PROs | | | | | |
| 2.45. | ITT | Eff_T16 | Summary of Responses to EQ-5D-5L Questions. | By Treatment Group and Visit<br>5 questions, 5 responses (no, slight, moderate, severe, unable to).<br>For V2, V3 and V6<br>Note: N, n, Mean, SD, Min, Max, Median | SAC |

| | | | | | |
|---|---|---|---|---|---|
| 2.46. | ITT | Eff_T1 | Summary of Derived EQ-5D-5L Utility Index Score by severity of Itch score at Baseline. | By Treatment Group and Visit<br>For V2, V3 and V6<br>Note: N, n, Mean, SD, Min, Max, Median | SAC |
| 2.47. | ITT | Eff_T1 | Summary of Change from Baseline in Derived EQ-5D-5L Utility Index Score by Severity Itch score at Baseline. | By Treatment Group and Visit<br>For V6<br>Note: N, n, Mean, SD, Min, Max, Median | SAC |
| 2.48. | ITT | Eff_T1 | Summary of EQ-5D-5L VAS. | By Treatment Group and Visit<br><br>For V2, V3 and V6<br><br>Note: N, n, Mean, SD, Min, Max, Median | SAC |
| 2.49. | ITT | Eff_T1 | Summary of Change from Baseline in EQ-5D-5L VAS. | For V6<br><br>Note: N, n, Mean, SD, Min, Max, Median | SAC |
| 2.50. | ITT | Eff_T4a | ANCOVA Analysis of Change from Baseline in Derived EQ-5D-5L Utility Index Score and EQ-5D-5L VAS at Week 16 (V6). | LS Mean Change from Baseline with 95% CI for both EQ-5D-5L utility index score and VAS | SAC |
| 2.51. | ITT | Eff_T16 | Summary of PGI-S. | PGI-S should be shown from V2 to V7 (inclusive). Show counts and percentages for each response. | SAC |
| 2.52. | ITT | Eff_T17 | Shift Table of Change from Baseline in PGI-S. | Show counts and percentages in each response pair (i.e. Severe at baseline, slight at Visit 4)<br>From V4 to V7 | SAC |
| 2.53. | ITT | Eff_T1 | Summary of Derived PGI-S score. | PGI-S should be shown from V2 to V7 (inclusive). | SAC |
| 2.54. | ITT | Eff_T1 | Summary of Change from Baseline in Derived PGI-S Score. | From V4 to V7 | SAC |
| 2.55. | ITT | Eff_T4a | ANCOVA Analysis of LS Mean Change from Baseline in PGI-S at Week 16 (V6) with 95% Confidence Intervals. | | SAC |
| 2.56. | ITT | Eff_T16 | Summary of PGI-C | PGI-C should be shown from V3 to V7 (inclusive). | SAC |
| 2.57. | ITT | Eff_T1 | Summary of Derived PGI-C Score. | PGI-C should be shown from V3 to V7 (inclusive). | SAC |
| 2.58. | ITT | Eff_T4a | ANCOVA Analysis of PGI-C at Week 16 (V6). | LS Mean PGI-C score (with 95% CI), Mean difference for each treatment group from Placebo, along with mean difference between 90 mg BID and 180 mg QD, for the derived PGI-C score (with 95% CI) | SAC |
| 2.59. | ITT | Eff_T1 | Summary of BDI-II Score. | For V1, V3 and V6<br>Note: N, n, Mean, SD, Min, Max, Median | SAC |

| | | | | | |
|---|---|---|---|---|---|
| 2.60. | ITT | Eff_T1 | Summary of Change from Baseline in BDI-II Score. | For V6 | SAC |
| 2.61. | ITT | Eff_T4a | ANCOVA Analysis of Change from Baseline in BDI-II Score at Week 16 (V6). | | SAC |
| 2.62. | ITT | Eff_T12 | Summary of Participant Treatment Experience Assessment (PTEA) on Treatment Benefits and Disadvantages at Week 16 (V6) | | HLR |
| **High Risk Population – Risk Responders** | | | | | |
| 2.63. | High Risk | Eff_T9 | Summary of Disease Progression Risk Responders. | For V2 to V7 (inclusive).<br>Footnotes:<br>Note: Participants were assigned to the High Risk Group at randomisation based on lab results at Visit 2.<br><br>Note: Risk Responder criteria is ALP < 1.67 * ULN *AND* total bilirubin ≤ ULN at the relevant Visit. | SAC |
| 2.64. | High Risk | Eff_T10 | Logistic Regression Analysis of Disease Progression Risk at Week 16 (V6). | Add footnotes:<br>Note: Participants were assigned to the High Risk Group at randomisation based on lab results at Visit 2.<br><br>Note: Disease Progression Risk Responder criteria is ALP at Week 16 (V6) < 1.67 * ULN *AND* total bilirubin at Week 16 (V6) ≤ ULN. | SAC |
| 2.65. | High Risk | Eff_T20 | ANCOVA Analysis of Percentage Change from Baseline in Serum ALP Concentrations at Week 16 (V6) | Include only placebo,180QD, and 90BID.<br>Use BMI as covariate in the Model to adjust for. | HLR |
| **High Risk Population - Markers of Disease** | | | | | |
| 2.66. | High Risk | LB1 | Summary of Markers of Disease | Markers of Disease are:<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR | SAC |
| 2.67. | High Risk | LB1 | Summary of Change from Baseline (V3) at Week 16 (V6) in Markers of Disease. | Markers of Disease are:<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR | HLR |
| 2.68. | High Risk | Eff_T4a | ANCOVA Analysis of Change from Baseline in Markers of Disease at Week 16 (V6). | This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups.<br>Markers of Disease are:<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR | SAC |

| | | | | | |
|---|---|---|---|---|---|
| **Restricted High Risk Population – Risk Responders** | | | | | |
| 2.69. | Restricted High Risk | Eff_T9 | Summary of Disease Progression Risk Responders. | For V2 to V7 (inclusive).<br>Footnotes:<br>Note: Those assigned to High Risk strata for randomisation who were still at High Risk at Baseline (confirmed as at High Risk at V2 and V3 based on lab results).<br><br>Note: Risk Responder criteria is ALP < 1.67 * ULN *AND* total bilirubin ≤ ULN at the relevant Visit. | SAC |
| 2.70. | Restricted High Risk | Eff_T10 | Logistic Regression Analysis of Disease Progression Risk at Week 16 (V6). | Add footnotes:<br>Note: Those assigned to High Risk strata for randomisation who were still at High Risk at Baseline (High Risk at V2 and V3 based on lab results).<br><br>Note: Disease Progression Risk Responder criteria is ALP at Week 16 (V6) < 1.67 * ULN *AND* total bilirubin at Week 16 (V6) ≤ ULN. | SAC |
| **Restricted High Risk Population - Markers of Disease** | | | | | |
| 2.71. | Restricted High Risk | LB1 | Summary of Change from Baseline (V3) at Week 16 (V6) in Markers of Disease | Note: Those assigned to High Risk strata for randomisation who were still at High Risk at Baseline (High Risk at V2 and V3 based on lab results).<br>Markers of Disease are:<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR | SAC |
| 2.72. | Restricted High Risk | Eff_T4a | ANCOVA Analysis of Change from Baseline in Markers of Disease at Week 16 (V6). | Note: Those assigned to High Risk strata for randomisation who were still at High Risk at Baseline (High Risk at V2 and V3 based on lab results).<br>This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups.<br>Markers of Disease are:<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR | SAC |

| | | | | | |
|---|---|---|---|---|---|
| **Country** | | | | | |
| **Mean Worst Daily Itch** | | | | | |
| 2.73. | ITT | Eff_T1 | Summary of Mean Worst Daily Itch by Country | For V3 to V7 (inclusive) | SAC |
| 2.74. | ITT | Eff_T1 | Summary of Change from Baseline in Mean Worst Daily Itch by Country | For V3 to V7 (inclusive) | SAC |
| **Itch Responders** | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| 2.75. | ITT | Eff_T9 | Summary of Mean Worst Daily Itch Responders by Country | The three definitions should be shown in the one table.<br>For V4 to V7 (inclusive) | SAC |
| **Responder Days** | | | | | |
| 2.76. | ITT | Eff_T1 | Summary of Responder Days in Mean Worst Daily Itch During the Main Study Period by Country | By treatment group and visit<br>Include all 3 responder definitions in one table with column added to left of table to identify the response definition<br>Note: Percentage of Responder Days in Mean Worst Daily Itch: number of days response V3+1 and V6-1/ number of days mean worst daily itch scores available | SAC |
| **Time to Worsening of Mean Worst Daily Itch** | | | | | |
| 2.77. | Responders (>=2 point improvement) | Eff_T1 | Summary of Time to Worsening of Itch following Visit 6 in those that have improved between Baseline and Visit 6 by Country | Footnotes as below:<br>Note: Improvement is defined as those that have a change from baseline in mean worst daily itch of at least -2.<br>Note: A participant is defined as worsened at the time their rolling average Mean Worst Daily Itch is at least 2 higher than their Mean Worst Daily Itch at Visit 6.<br><br>Note: Treatment groups are only included in this analysis if they consist of at least 8 participants. | SAC |
| **Other eDiary itch questions** | | | | | |
| 2.78. | ITT | Eff_T1 | Summary of Mean Daily Sleep Score by Country | Footnote:<br>Note: Mean Daily Sleep Score is defined as the average of the daily sleep scores provided in the 7 days prior to the relevant visit. | SAC |
| 2.79. | ITT | Eff_T1 | Summary of Change from Baseline in Mean Daily Sleep Score by Country | Note: Mean Daily Sleep Score is defined as the average of the daily sleep scores provided in the 7 days prior to the relevant visit. | SAC |
| 2.80. | ITT | Eff_T1 | Summary of Mean Daily Fatigue Score by Country | Note: Mean Daily Fatigue Score is defined as the average of the daily fatigue scores provided in the 7 days prior to the relevant visit. | SAC |
| 2.81. | ITT | Eff_T1 | Summary of Change from Baseline in Mean Daily Fatigue Score by Country | Note: Mean Daily Fatigue Score is defined as the average of the daily fatigue scores provided in the 7 days prior to the relevant visit. | SAC |

| | | | | | |
|---|---|---|---|---|---|
| 2.82. | ITT | Eff_T1 | Summary of Mean Overall Itch Score by Country | Footnote: Note: Mean Overall Itching Score is defined as the average of the mean of the morning and evening overall itching scores provided in the 7 days prior to the relevant visit. | SAC |
| 2.83. | ITT | Eff_T1 | Summary of Change from Baseline in Mean Overall Itch Score by Country | Note: Mean Overall Itch Score is defined as the average of the mean of the morning and evening overall itch scores provided in the 7 days prior to the relevant visit. | SAC |
| 2.84. | ITT | Eff_T13 | Summary of Average Response to Categorical Itch Questions in eDiary (AM) by Country | | SAC |
| 2.85. | ITT | Eff_T13 | Summary of Most Frequent Response for Categorical Itch Questions in eDiary (AM) by Country | Repeat, paginated by question with full question in sub-title, for all questions in e-diary (AM) that are not on 0-10 NRS.<br>Footnote: Most Frequent Response is taken as the mode of responses for each participant in the 7 days prior to each Visit.<br>Footnote: If there is more than one mode the Most Frequent Response is taken to be the worst of the modes. | SAC |
| 2.86. | ITT | Eff_T13 | Summary of Average Response to Categorical Itch Questions in eDiary (PM) by Country | Repeat, paginated by question with full question in sub-title, for all questions in e-diary (PM) that are not on 0-10 NRS. | SAC |
| 2.87. | ITT | Eff_T13 | Summary of Most Frequent Response to Categorical Itch Questions in eDiary (PM) by Country | Repeat, paginated by question with full question in sub-title, for all questions in e-diary (PM) that are not on 0-10 NRS.<br>Footnote: Most Frequent Response is taken as the mode of responses for each participant in the 7 days prior to each Visit.<br>Footnote: If there is more than one mode the Most Frequent Response is taken to be the worst of the modes. | SAC |
| 2.88. | ITT | Eff_T1 | Summary of Itch, Fatigue and Sleep at Follow-up (Visit 8) by Country | Table should include a summary of the scores give to worst itching, itching interfere with sleep and worst fatigue at the follow-up telephone call. | SAC |
| 2.89. | ITT | Eff_T13 | Summary of Response to Categorical Itch Questions at Follow-up (Visit 8) by Country | Table should include counts and percentages of the responses to the categorical questions asked during the follow-up phone call | SAC |

| | | | | | |
|---|---|---|---|---|---|
| Validated Questionnaires | | | | | |
| 5-D Itch | | | | | |
| 2.90. | ITT | Eff_T1 | Summary of 5-D Itch Scores. by Country | By treatment group and visit<br>Note: N, n, Mean, SD, Min, Max, Median<br>Paginated by domain, with domain (as below) in subtitle. Only for V2, V3, V6 and V7.<br>Duration, Degree, Direction, Disability, Distribution, Total 5-D Itch Score | SAC |
| 2.91. | ITT | Eff_T1 | Summary of Change from Baseline in 5-D Itch Scores by Country | Paginated by domain, with domain (as below) in subtitle. Only for, V6 and V7.<br>Duration, Degree, Direction, Disability, Distribution, Total 5-D Itch Score | SAC |
| PBC-40 | | | | | |
| 2.92. | ITT | Eff_T1 | Summary of PBC-40 Scores by Country | Paginated by domain, with domain (as below) in subtitle.<br>Itch, Fatigue, Symptoms, Cognitive, Social, Emotional<br>For V1 to V7 (inclusive)<br><br>Note: Summary (N, n, Mean, SD, Min, Max, Median) by Treatment Group | SAC |
| 2.93. | ITT | Eff_T1 | Summary of Change from Baseline in PBC-40 Scores by Country | Paginated by domain, with domain (as below) in subtitle.<br>Itch, Fatigue, Symptoms, Cognitive, Social, Emotional<br>For V4 to V7 (inclusive)<br>Note: Summary (N, n, Mean, SD, Min, Max, Median) by Treatment Group | SAC |
| Other PROs | | | | | |
| 2.94. | ITT | Eff_T16 | Summary of Responses to EQ-5D-5L Questions by Country | By Treatment Group and Visit<br>5 questions, 5 responses (no, slight, moderate, severe, unable to).<br>For V2, V3 and V6<br>Note: N, n, Mean ,SD, Min, Max, Median | SAC |
| 2.95. | ITT | Eff_T1 | Summary of Derived EQ-5D-5L Utility Index Score by severity of Itch score at Baseline by Country | By Treatment Group and Visit<br>For V2, V3 and V6<br>Note: N, n, Mean ,SD, Min, Max, Median | SAC |
| 2.96. | ITT | Eff_T1 | Summary of Change from Baseline in Derived EQ-5D-5L Utility Index Score by severity of Itch score at Baseline by Country | By Treatment Group and Visit<br>For V6<br>Note: N, n, Mean ,SD, Min, Max, Median | SAC |

| | | | | | |
|---|---|---|---|---|---|
| 2.97. | ITT | Eff_T1 | Summary of EQ-5D-5L VAS by Country | By Treatment Group and Visit<br>For V2, V3 and V6<br>Note: N, n, Mean ,SD, Min, Max, Median | SAC |
| 2.98. | ITT | Eff_T1 | Summary of Change from Baseline in EQ-5D-5L VAS by Country | For V6<br>Note: N, n, Mean ,SD, Min, Max, Median | SAC |
| 2.99. | ITT | Eff_T16 | Summary of PGI-S by Country | PGI-S should be shown from V2 to V7 (inclusive). Show counts and percentages for each response. | SAC |
| 2.100. | ITT | Eff_T17 | Shift Table of Change from Baseline in PGI-S by Country | Show counts and percentages in each response pair (i.e. Severe at baseline, slight at Visit 4)<br>From V4 to V7 | SAC |
| 2.101. | ITT | Eff_T1 | Summary of Derived PGI-S score by Country | PGI-S should be shown from V2 to V7 (inclusive). | SAC |
| 2.102. | ITT | Eff_T1 | Summary of Change from Baseline in Derived PGI-S Score by Country | From V4 to V7 | SAC |
| 2.103. | ITT | Eff_T16 | Summary of PGI-C by Country | PGI-C should be shown from V3 to V7 (inclusive). | SAC |
| 2.104. | ITT | Eff_T1 | Summary of Derived PGI-C Score by Country | PGI-C should be shown from V3 to V7 (inclusive). | SAC |
| 2.105. | ITT | Eff_T1 | Summary of BDI-II Score by Country | For V1, V3 and V6<br>Note: N, n, Mean, SD, Min, Max, Median | SAC |
| 2.106. | ITT | Eff_T1 | Summary of Change from Baseline in BDI-II Score by Country | For V6 | SAC |
| 2.107. | ITT | Eff_T12 | Summary of Participant Treatment Experience Assessment (PTEA) on Treatment Benefits and Disadvantages at Week 16 (V6) by Country | | SAC |
| **High Risk Population – Risk Responders** | | | | | |
| 2.108. | High Risk | Eff_T9 | Summary of Disease Progression Risk Responders by Country | For V2 to V7 (inclusive).<br>Footnotes:<br>Note: Participants were assigned to the High Risk Group at randomisation based on lab results at Visit 2.<br>Note: Risk Responder criteria is ALP < 1.67 * ULN *AND* total bilirubin ≤ ULN at the relevant Visit. | SAC |

| | | | | | |
|---|---|---|---|---|---|
| **High Risk Population - Markers of Disease** | | | | | |
| 2.109. | High Risk | LB1 | Summary of Markers of Disease by Country | Markers of Disease are:<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR | SAC |
| 2.110. | High Risk | LB1 | Summary of Change from Baseline (V3) at Week 16 (V6) in Markers of Disease by Country | Markers of Disease are:<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR | SAC |

### 12.12.6. Efficacy Figures

| Efficacy: Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Mean Worst Daily Itch** | | | | | |
| 2.1. | ITT | Eff_F18 | Plot of Rolling Average Worst Daily Itch Score against Study Day by Treatment Group | From Visit 2 (Day 1) to Visit 7 (end of final study period).<br><br>Note:<br>The rolling average is calculated as the average worst daily itch during the 7 days prior to each study day, within each study period. | SAC |
| 2.2. | ITT | Eff_F1 | Individual Mean Worst Daily Itch against Day: with AM and PM Itch for 7 Days Prior to Each Visit (V3 to V7) by Treatment Group | with Participant ID in plot subtitle not in the title.<br>Scatterplot of 7 days scores (AM and PM) prior to visit, with different symbols/colours for AM and PM.<br>Overlay the mean worst daily itch for each Visit, spanning the 7 days that contribute to it on the plot.<br>Add ref line at the day of each Visit and label it with appropriate visit number<br>This should include Visits 3 to 7 inclusive. | SAC |
| 2.3. | ITT | Eff_F2 | Spaghetti Plot of Change from Baseline in Mean Worst Daily Itch against Visit: Panelled by Treatment Group and Risk Group. | Do not include legend | SAC |
| 2.4. | ITT | Eff_F3 | Boxplot of Change from Baseline in Mean Worst Daily Itch grouped by Treatment Group: Panelled by Visit | Visits 4-7 inclusive. Include ref line at y=0 | SAC |
| 2.5. | ITT | Eff_F4 | Line Plot of Mean Treatment Group Change from Baseline in Mean Worst Daily Itch against Visit | Line for each treatment group. Include markers to differentiate BID and QD dosing. Visits 4-7 inclusive. Include ref line at y=0 | SAC |
| 2.6. | ITT | Eff_F4 | Line Plot of Mean Treatment Group Percent Change from Baseline Mean Worst Daily Itch against Visit | Line for each treatment group. Include markers to differentiate BID and QD dosing. Visits 4-7 inclusive. Include ref line at y=0 | SAC |
| 2.7. | ITT | Eff_F5 | Preliminary Checks: ANCOVA: Estimated Mean Difference from Placebo in Change from Baseline in Mean Worst Daily Itch at Week 16 (V6) with 95% Confidence Intervals. | Include ref line at y=0 | SAC |

| Efficacy: Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 2.8. | ITT | Eff_F6 | Preliminary Checks: Linear Model: Estimated Mean Difference from Placebo in Change from Baseline in Mean Worst Daily Itch at Week 16 (V6) with 95% Confidence Intervals Against Total Daily Dose. | Include ref line at y=0 | SAC |
| 2.9. | ITT | Eff_F6 | See Appendix 14 | | SAC |
| 2.10. | ITT | Eff_F7a or Eff_F7b | Estimated Mean Difference from Placebo in Change from Baseline in Mean Worst Daily Itch at Week 16 (V6) with 95% Confidence Intervals Against Total Daily Dose. | If the final chosen model is ANCOVA then the x-axis will be categorical and will include the treatment groups (as per Eff_F7a)<br>Alternatively the x-axis will be continuous and will include the total daily dose levels used in the study (as per Eff_F7b).<br>Include ref line at y=0 | SAC |
| 2.11. | ITT | Eff_F8 | Fitted Dose-Response Model for Change from Baseline in Mean Worst Daily Itch with 95% Prediction Intervals Against Total Daily Dose. | Include ref line at y=0<br>Change from baseline at Week 16 will be overlaid as a scatter plot. | SAC |
| 2.12. | ITT | Eff_F8 | Fitted Dose-Response Model for Change from Baseline in Mean Worst Daily Itch with 95% Prediction Intervals Against Total Daily Dose (QD Dose levels only). | Include ref line at y=0<br>Change from baseline at Week 16 will be overlaid as a scatter plot QD Dose Levels only) | SAC |
| 2.13. | ITT | Eff_F8 | Fitted Dose-Response Model for Change from Baseline in Mean Worst Daily Itch with 95% Prediction Intervals Against Total Daily Dose (BID Dose Levels only). | Include ref line at y=0<br>Change from baseline at Week 16 will be overlaid as a scatter plot BID Dose Levels only. | SAC |
| 2.14. | ITT | Eff_F8 | Fitted Dose-Response Model for Change from Baseline in Mean Worst Daily Itch with 95% Prediction Intervals Against Total Daily Dose (overlay of QD and BID Dose). | Include ref line at y=0<br>Change from baseline at Week 16 will be overlaid as a scatter plot Overlay of QD and BID Dose levels on one figure (2 curves overlay). | HLR |
| 2.15. | ITT | Eff_F8 | Fitted Dose-Response Model for Change from Baseline in Mean Worst Daily Itch with 95% Prediction Intervals Against Total Daily Dose (Convert 180 QD Dose level to 90 BID Dose level). | Include ref line at y=0<br>Change from baseline at Week 16 will be overlaid as a scatter plot convert 180 QD dose level to 90 BID does level. | SAC |
| 2.16. | ITT | Eff_F12 | MMRM: Mean Difference from Placebo in Change from Baseline in Mean Worst Daily Itch at Week 16 (V6) with 95% Confidence Intervals | Include ref line at y=0 | SAC |
| 2.17. | ITT | Eff_F17 | Estimated bias associated with ignoring the impact of Risk Group | ONLY if Risk Group is significant in ANCOVA | SAC |

| Efficacy: Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 2.18. | ITT | Eff_F20 | Pie Chart of Baseline Itch category and at Week 16 (V6) by Treatment Group | .<br><br>Use itch category as by variables Repeat figure by treatment group<br><br>Programming notes: please add n (%), see shell for more details | HLR |
| 2.19. | ITT | Eff_F22 | Line Plot of Mean ALP over the Main Study Period by Treatment Group | Add a reference line at 125 and (N=) in addition to the treatment group in the legend<br>x-axis is visit in weeks. | SAC |
| **Itch Responders** | | | | | |
| 2.20. | ITT | Eff_F14 | Odds Ratio of Responder in Mean Worst Daily Itch at Week 16 (V6), vs. Placebo, with 95% Confidence Intervals. | Panelled by response definition | SAC |
| **Itch Responder Days** | | | | | |
| 2.21. | ITT | Eff_F15 | Boxplot of Percentage of Responder Days in Mean Worst Daily Itch over Main Study Period. | Include each of the 3 response definitions on one figure.<br>Panelled by responder definition | SAC |
| 2.22. | ITT | Eff_F7a | Mean Difference from placebo in Percentage of Responder Days in Mean Worst Daily Itch over Main Study Period with 95% Confidence Intervals. | Panelled by Responder definition | SAC |
| **Time to Worsening of Mean Worst Daily Itch** | | | | | |
| 2.23. | Responders (>=2 point improvement)s | TTE10 | Kaplan-Meier Plot of Time to Worsening of Mean Worst Daily Itch following Visit 6. | | SAC |
| **Other eDiary itch questions** | | | | | |
| 2.24. | ITT | Eff_F4 | Line Plot of Mean Sleep, Fatigue and Overall Itch Scores (NRS) against Visit for each Treatment Group | Line for each treatment group. Include markers.<br>Visits 3-7 inclusive<br>Panelled by endpoint – one plot each for Mean Sleep, mean fatigue and mean overall itch within one page. | SAC |
| 2.25. | ITT | Eff_F7a | Mean Difference from placebo in Change from Baseline in Sleep, Fatigue and Overall Itch Scores (NRS) at Week 16 (V6) with 95% Confidence Intervals. | Change from baseline.<br>Include Mean Sleep Score, Mean Fatigue Score and Mean Overall Itch on one page – panelled by endpoint | SAC |

| Efficacy: Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Validated Questionnaires** | | | | | |
| **5-D Itch** | | | | | |
| 2.26. | ITT | Eff_F16 | Boxplot of Change from Baseline in 5-D Itch Scores against Visit: Panelled by Domain | Panelled by 5 5-D Itch domains and total score in a 2*3 grid (plot each Domain per page, since the scale differ per domain) | SAC |
| 2.27. | ITT | Eff_7a | Mean Difference from placebo in Change from Baseline in 5-D Itch at Week 16 (V6), with 95% Confidence Intervals. | Change from baseline.<br>Panelled by 5-D Itch domain and Total score in a 3*2 grid | SAC |
| **PBC-40** | | | | | |
| 2.28. | ITT | Eff_F16 | Boxplot of PBC-40 Scores against Visit: Panelled by Domain | Panelled by PBC-40 domain | SAC |
| 2.29. | ITT | Eff_F7a | Mean Difference from Placebo in Change from Baseline in PBC-40 Score at Week 16 (V6), with 95% Confidence Intervals. | Change from baseline.<br>Panelled by PBC-40 domain | SAC |
| **Other PROs** | | | | | |
| 2.30. | ITT | EFF_F4 | Line Plot of Derived EQ-5D-5L Utility Index Score against Visit. | Different line for each treatment group | SAC |
| 2.31. | ITT | Eff_F4 | Line Plot of EQ-5D-VAS against Visit. | Different line for each treatment group | SAC |
| 2.32. | ITT | Eff_F7a | Mean Difference from Placebo in Change from Baseline in Derived EQ-5D-5L Utility Index Score and EQ-5D-5L VAS at Week 16 (V6), with 95% Confidence Intervals. | Change from baseline.<br>Panel for each of the two scores | SAC |
| 2.33. | ITT | Eff_F16 | Boxplot of PGI-S Scores by Visit and Treatment Group | | SAC |
| 2.34. | ITT | Eff_F13 | Bar Chart of PGI-C Response against Visit | Percentage of Participants in each treatment group with each response – with a panel for each visit the score is available | SAC |
| 2.35. | ITT | Eff_F16 | Boxplot of PGI-C Scores by Visit and Treatment Group | | SAC |
| 2.36. | ITT | Eff_F16 | Boxplot of BDI-II Scores against Visit | | SAC |
| 2.37. | ITT | Eff_F7a | Mean Difference from Placebo in Change from Baseline in BDI-II at Week 16 (V6), with 95% Confidence Intervals. | Change from baseline. | SAC |
| **High Risk Population – Risk Response** | | | | | |
| 2.38. | High Risk | Eff_F13 | Bar chart of Percentage of Responders in Disease Progression Risk at each Visit). | Percentage of Responders.<br>Y-axis range from 0-100%. | SAC |
| 2.39. | High Risk | Eff_F14 | Odds Ratio of Response in Disease Progression Risk at Week 16 (V6), vs. Placebo, with 95% Confidence Intervals. | | SAC |

| Efficacy: Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 2.40. | High Risk | LB11 | Line Plot of Markers of Disease against Visit. | One page per patient – panelled plot with 3 plots<br>1) ALP, AST, ALT,<br>2) GGT, Albumin, Total Bilirubin–<br>3) PT/INR<br><br>Plot 1 and 2 should use a different colour for each marker (should be set up in attribute map so colours are consistent across all Participants)<br>The y-axis for plot 1 and 2 should be marker/ULN, may be more appropriate for this axis to be displayed on the log-scale if it makes the data display clearer.<br>The levels of clinical concern for each marker should be shown on the plot | SAC |
| 2.41. | High Risk | LB9 | Boxplots of Markers of Disease against Visit. | Boxplot of each marker of disease for each visit.<br>Panel by Treatment Group<br>One page per Marker<br>Different colour for treatment group (set up an attribute map so colours are consistent)<br>The y-axis will be the marker/ULN, may be more appropriate for this axis to be displayed on the log-scale if it makes the data display clearer.<br>The levels of clinical concern should be shown as a line on the plot | SAC |
| 2.42. | High Risk | Eff_F7a | Mean Difference from Placebo in Change from Baseline in Markers of Disease at Week 16 (V6) with 95% Confidence Intervals. | Change from baseline.<br>Panelled by marker | SAC |

| | | | | | |
|---|---|---|---|---|---|
| **Country** | | | | | |
| 2.43. | ITT | Eff_F19a or Eff_F19b | Estimated Mean Difference from Placebo in Change from Baseline in Mean Worst Daily Itch at Week 16 (V6) with 95% Confidence Intervals Against Total Daily Dose by Country. | If the final chosen model is ANCOVA then the x-axis will be categorical and will include the treatment groups (as per Eff_F19a)<br>Alternatively the x-axis will be continuous and will include the total daily dose levels used in the study (as per Eff_F19b).<br>Include ref line at y=0.<br>Programming notes: Change symbols to different shapes | SAC |
| **Itch Responders** | | | | | |
| 2.44. | ITT | Eff_F19a | Odds Ratio of Responder in Mean Worst Daily Itch at Week 16 (V6), vs. Placebo, with 95% Confidence Intervals by Country. | Panelled by response definition. Eff_F19a but for odds ratio<br>Programming notes: Change symbols to different shapes | SAC |
| **Itch Responder Days** | | | | | |
| 2.45. | ITT | Eff_F19a | Mean Difference from placebo in Percentage of Responder Days in Mean Worst Daily Itch over Main Study Period with 95% Confidence Intervals by Country. | Panelled by Responder definition<br>Programming notes: Change symbols to different shapes | SAC |
| **Other eDiary itch questions** | | | | | |
| 2.46. | ITT | Eff_F19a | Mean Difference from placebo in Change from Baseline in Sleep, Fatigue and Overall Itch at Week 16 (V6) with 95% Confidence Intervals by Country. | Change from baseline.<br>Include Mean Sleep Score, Mean Fatigue Score and Mean Overall Itch on one page – panelled by endpoint<br>Programming notes: Change symbols to different shapes | SAC |

| | | | | | |
|---|---|---|---|---|---|
| **Validated Questionnaires** | | | | | |
| **5-D Itch** | | | | | |
| 2.47. | ITT | Eff_19a | Mean Difference from placebo in Change from Baseline in 5-D Itch at Week 16 (V6), with 95% Confidence Intervals by Country. | Change from baseline.<br>Panelled by 5-D Itch domain and Total score in a 3*2 grid<br>Programming notes: Change symbols to different shapes | SAC |
| **PBC-40** | | | | | |
| 2.48. | ITT | Eff_F19a | Mean Difference from Placebo in Change from Baseline in PBC-40 Score at Week 16 (V6), with 95% Confidence Intervals by Country. | Change from baseline.<br>Panelled by PBC-40 domain<br>Programming notes: Change symbols to different shapes | SAC |
| **Other PROs** | | | | | |
| 2.49. | ITT | Eff_F19a | Mean Difference from Placebo in Change from Baseline in Derived EQ-5D-5L Utility Index Score and EQ-5D-5L VAS at Week 16 (V6), with 95% Confidence Intervals by Country. | Change from baseline.<br>Panel for each of the two scores<br>Programming notes: Change symbols to different shapes | SAC |
| 2.50. | ITT | Eff_F19a | Mean Difference from Placebo in Change from Baseline in PGI-S at Week 16 (V6), with 95% Confidence Intervals by Country. | Programming notes: Change symbols to different shapes | SAC |
| 2.51. | ITT | Eff_F19a | Mean difference from Placebo in Change from Baseline in PGI-C at Week 16 (V6), with 95% Confidence Intervals by Country. | Not change from baseline.<br>Programming notes: Change symbols to different shapes | SAC |
| 2.52. | ITT | Eff_F19a | Mean Difference from Placebo in Change from Baseline in BDI-II at Week 16 (V6), with 95% Confidence Intervals by Country. | Change from baseline.<br>Programming notes: Change symbols to different shapes | SAC |
| **High Risk Population – Risk Response** | | | | | |
| 2.53. | High Risk | Eff_F19a | Mean Difference from Placebo in Change from Baseline in Markers of Disease at Week 16 (V6) with 95% Confidence Intervals by Country. | Change from baseline.<br>Panelled by marker<br>Programming notes: Change symbols to different shapes | SAC |

### 12.12.7. Safety Tables

| Safety : Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Adverse Events** | | | | | |
| 3.1. | Safety | AE1 | Summary of All Adverse Events | Number of adverse events by SOC and PT | HLR |
| 3.2. | Initial | AE1 | Summary of Pre-Treatment Adverse Events | Footnote: Pre-Treatment is defined as the period from start of Initial Study Period (V2) until early withdrawal of study treatment or V3, whichever is earliest. | SAC |
| 3.3. | Safety | AE1 | Summary of On-Treatment Adverse Events | Footnote: On-Treatment is defined as from the date if first randomised study treatment until the date of last randomised study treatment + 2 days | SAC |
| 3.4. | Safety | AE1 | Summary of Post-Treatment Adverse Events | Footnote: Note: Post-Treatment is defined as within 30 days of date of last randomised study treatment<br>Programming note:Post-treatment is defined as date of last randomised study treatment + 30 days. | SAC |
| 3.5. | Safety | AE1 | Summary of On-Treatment Drug-Related Adverse Events | | SAC |
| 3.6. | Safety | AE3 | Summary of Common (>=5%) Adverse Events by Overall Frequency | | SAC |
| 3.7. | Safety | AE3 | Summary of the 10 Most Frequent On-treatment Adverse Events in Each Treatment Group | Count most frequent number of events, not Participants and %. Sort by descending total incidence (sum of number of Participants with event). | SAC |
| 3.8. | Initial | AE1 | Summary of Pre-Treatment Fatal Serious Adverse Events | | SAC |
| 3.9. | Safety | AE1 | Summary of On-Treatment Fatal Serious Adverse Events | | SAC |
| 3.10. | Safety | AE1 | Summary of Post-Treatment Fatal Serious Adverse Events | | SAC |
| 3.11. | Safety | AE1 | Summary of On-Treatment Drug-Related Fatal Serious Adverse Events | | SAC |
| 3.12. | Initial | AE1 | Summary of Pre-Treatment Non-Fatal Serious Adverse Events | | SAC |

| Safety : Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 3.13. | Safety | AE1 | Summary of On-Treatment Non-Fatal Serious Adverse Events | | SAC |
| 3.14. | Safety | AE1 | Summary of Post-Treatment Non-Fatal Serious Adverse Events | | SAC |
| 3.15. | Safety | AE1 | Summary of On-Treatment Drug-Related Non-Fatal Serious Adverse Events | | SAC |
| 3.16. | Safety | AE1 | Summary of Adverse Events Leading to Permanent Discontinuation of Study Drug | Note: use the main study period part of this table for HLR for presentation purpose only (no change to the standard table). | HLR |
| 3.17. | Safety | AE1 | Summary of Adverse Events Leading to Permanent Discontinuation of Study Drug (during placebo run-in) | | SAC |
| 3.18. | Safety | AE1 | Summary of Adverse Events Leading to Withdrawal from Study | | SAC |
| 3.19. | Safety | AE1 | Summary of Adverse Events Leading to Withdrawal from Study (during placebo run-in) | | SAC |
| 3.20. | Initial | AE1 | Summary of Pre-Treatment Diarrhoea Adverse Events of Special Interest | Diarrhoea adverse event as defined in RAP | SAC |
| 3.21. | Safety | AE1 | Summary of On-Treatment Diarrhoea Adverse Events of Special Interest | Diarrhoea adverse event as defined in RAP | HLR |
| 3.22. | Safety | AE1 | Summary of Post-Treatment Diarrhoea Adverse Events of Special Interest | Diarrhoea adverse event as defined in RAP | SAC |
| 3.23. | Safety | TTE7 | Time to Event Analysis for Time to Diarrhoea. | Diarrhoea adverse event as defined in RAP | SAC |
| 3.24. | Safety | AE1 | Summary of On-Treatment Abdominal Pain Adverse Events | Abdominal pain adverse event as defined in RAP | SAC |
| 3.25. | Safety | AE1 | Summary of Post-Treatment Abdominal Pain Adverse Events | Abdominal pain adverse event as defined in RAP | SAC |
| 3.26. | Safety | TTE7 | Time to Event Analysis for Time to Abdominal Pain. | Abdominal Pain adverse event as defined in RAP | SAC |
| 3.27. | Safety | AE1 | Number and Percent of Participants with Diarrhoea and Abdominal Pain during the Main Study Phase | | SAC |
| 3.28. | Initial | AE1 | Summary of Pre-Treatment Liver Adverse Events of Special Interest<br><br>Liver AESI Met the Stopping Criteria | ALT AESI as defined in RAP | SAC |

| Safety : Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 3.29. | Safety | AE1 | Summary of On-Treatment Liver Adverse Events of Special Interest<br><br>Liver AESI Met the Stopping Criteria | ALT AESI as defined in RAP | SAC |
| 3.30. | Safety | AE1 | Summary of Post-Treatment Liver Adverse Events of Special Interest<br><br>Liver AESI Met the Stopping Criteria | ALT AESI as defined in RAP | SAC |
| **Clinical Laboratory Parameters** | | | | | |
| **Chemistry Data** | | | | | |
| 3.31. | Safety | LB1 | Summary of Change from Baseline in Chemistry Data | Include absolute values for V1, V2, and V3 (Baseline) as well as change from baseline for all other visits<br>Chemistry Data includes:<br>BUN, creatinine, eGRF (CKD-EPI), Potassium, Sodium, Calcium. The liver function tests will be presented separately | SAC |
| 3.32. | Safety | LB17 | Summary of Chemistry Data of Potential Clinical Importance | Chemistry Data includes:<br>BUN, creatinine, eGRF (CKD-EPI), Potassium, Sodium, Calcium. The liver function tests will be presented separately | SAC |
| **Liver Function Tests** | | | | | |
| 3.33. | Safety | LB1 | Summary of Change from Baseline in Liver Function Tests | Include absolute values for V1, V2, and V3 (Baseline) as well as change from baseline for all other visits<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR. This will include central lab data only | SAC |

| Safety : Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 3.34. | Safety | LB1 | Summary of Change from Baseline in Liver Function Tests by Visit by Risk of Disease Progression | Include absolute values for V1, V2, and V3 (Baseline) as well as change from baseline at all other visits<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR. This will include central lab data only<br>(Note to GSK: For HLR slides presentation Pull out ALT, ALP, and Bilirubin only from this table) | HLR |
| 3.35. | Safety | LB1 | Summary of Ratio of Upper Limit of Normal in Liver Function Tests (Scheduled) | Include V1 – V7<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR This will include central lab data only | SAC |
| 3.36. | Safety | LB1 | Summary of Ratio of Upper Limit of Normal in Liver Function Tests (Unscheduled) | Include V1 – V7<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR This will include local lab data only and will present the data from unscheduled liver monitoring visits. | SAC |
| 3.37. | Safety | LB15 | Summary of Liver Function Tests: Normal Range Categories | This will include central lab data only | SAC |
| 3.38. | Safety | LB4 | Shift Table of Change from Baseline in Normal Range Categories for Liver Function Tests | Show proportion moving between the categories from baseline category. This will include central lab data only | SAC |
| 3.39. | Safety | LB17 | Summary of Change from Baseline in Liver Function Tests of Potential Clinical Importance | ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR. This will include central lab data only | SAC |
| 3.40. | Safety | Eff_T2 | ANCOVA Analysis of Change from Baseline in Liver Function Tests at Week 16 (V6). | ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR. This will include central lab data only | SAC |
| 3.41. | Safety | LB17 | Summary of Change from Baseline in Laboratory Data of Potential Clinical Importance | This will include all lab data with a change from baseline PCI flag – albumin, creatinine, haematocrit, haemoglobin, | SAC |

| Safety : Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Fasting Lipid Profile** | | | | | |
| **Vitamin Absorption** | | | | | |
| 3.42. | Safety | LB1 | Summary of Change from Baseline in Vitamin Absorption | Include absolute values for both V2 and V3 (Baseline) as well as change from baseline for all other visits<br>Vitamin Absorption includes:<br>Vitamin A, D, E and K | SAC |
| 3.43. | Safety | Eff_T2 | ANCOVA Analysis of Change from Baseline in Vitamin Absorption at Week 16 (V6). | This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups.<br>Vitamin Absorption includes:<br>Vitamin A, D, E and K | SAC |
| **Haematology** | | | | | |
| 3.44. | Safety | LB1 | Summary of Change from Baseline in Haematology Data | Include absolute values for both V2 and V3 (Baseline) as well as change from baseline for all other visits | SAC |
| 3.45. | Safety | LB17 | Summary of Haematology Data of Potential Clinical Importance | | SAC |
| **FOBT** | | | | | |
| 3.46. | Safety | LB1 | Summary of Faecal Occult Blood Data | | SAC |
| **Urinalysis** | | | | | |
| 3.47. | Safety | UR1 | Summary of Urinalysis Results by Visit | | SAC |
| **Hepatobiliary (Liver)** | | | | | |
| 3.48. | Safety | LIVER1 | Summary of Liver Monitoring/Stopping Event Reporting During Main Study Period | During MSP | HLR |
| 3.49. | Safety | LIVER1 | Summary of Liver Monitoring/Stopping Event Reporting During Initial Study Period | During ISP | SAC |

| Safety : Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Vital Signs and ECGs** | | | | | |
| 3.50. | Safety | VS1 | Summary of Change from Baseline in Vital Signs by Visit | Include absolute values for V1, V2, and V3 (Baseline) as well as change from baseline for all other visits. | SAC |
| 3.51. | Safety | EG1 | Summary of ECG Findings | Summary (freq and %) of Participants with normal and abnormal ECG results at each visit | SAC |
| 3.52. | Safety | EG2 | Summary of Change from Baseline in ECG Values by Visit | Include absolute values for V1, V2, and V3 (Baseline) as well as change from baseline for all other visits. | SAC |
| **GSRS** | | | | | |
| 3.53. | Safety | Eff_T16 | Summary of responses to GSRS questions | | SAC |
| 3.54. | Safety | Eff_T1 | Summary of GSRS Scores | As per Eff_T1, paginated by domain.<br><br>For V3 to V7 (inclusive) | SAC |
| 3.55. | Safety | Eff_T1 | Summary of Change from Baseline in GSRS Scores | As per Eff_T1, paginated by domain.<br><br>For V4 to V7 (inclusive) | SAC |
| 3.56. | Safety | EFF_T16 | Summary of Weekly e-diary GI symptoms questions (categorical) | The number and proportion of people choosing each of the categories for each of the symptoms questions.<br><br>As per Eff_T9 with a sub-table per question. | SAC |
| 3.57. | Safety | Eff_T1 | Summary of Weekly e-diary GI symptoms questions (0-10 NRS) | | SAC |
| 3.58. | Safety | Eff_T1 | Summary of Change from Baseline in Weekly e-diary GI symptoms questions (0-10 NRS) | | SAC |
| | Safety | AE1 | Summary of all Adverse Events by Country | Number of adverse events by SOC and PT | SAC |

| Safety : Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 3.60. | Safety | AE1 | Summary of On-Treatment Drug-Related Adverse Events by Country | | SAC |
| 3.61. | Safety | AE1 | Summary of Adverse Events Leading to Permanent Discontinuation of Study Drug by Country | | SAC |
| 3.62. | Safety | AE1 | Summary of Adverse Events Leading to Permanent Discontinuation of Study Drug (by Country, during placebo run-in) | | SAC |
| 3.63. | Safety | AE1 | Summary of Adverse Events Leading to Withdrawal from Study by Country | | SAC |
| 3.64. | Safety | AE1 | Summary of Adverse Events Leading to Withdrawal from Study (by Country, during placebo run-in) | | SAC |
| 3.65. | Safety | AE1 | Summary of On-Treatment Diarrhoea Adverse Events of Special Interest, by Country | Diarrhoea adverse event as defined in RAP | SAC |
| 3.66. | Safety | AE1 | Summary of On-Treatment Abdominal Pain Adverse Events by Country | Abdominal Pain adverse event as defined in RAP | SAC |
| 3.67. | Safety | AE1 | Summary of On-Treatment Liver Adverse Events of Special Interest by Country | ALT AESI as defined in RAP | SAC |
| **Clinical Laboratory Parameters** | | | | | |
| **Chemistry Data** | | | | | |
| 3.68. | Safety | LB1 | Summary of Change from Baseline in Chemistry Data by Country | Include absolute values for V1, V2, and V3 (Baseline) as well as change from baseline for all other visits<br>Chemistry Data includes:<br>BUN, creatinine, eGRF (CKD-EPI), Potassium, Sodium, Calcium. The liver function tests will be presented seperately | SAC |
| 3.69. | Safety | LB17 | Summary of Chemistry Data of Potential Clinical Importance by Country | Chemistry Data includes:<br>BUN, creatinine, eGRF (CKD-EPI), Potassium, Sodium, Calcium. The liver function tests will be presented seperately | SAC |

| Safety : Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Liver Function Tests** | | | | | |
| 3.70. | Safety | LB1 | Summary of Change from Baseline in Liver Function Tests by Country | Include absolute values for V1, V2 and V3 (Baseline) as well as change from baseline for all other visits<br>ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR. This will include central lab data only | SAC |
| 3.71. | Safety | LB15 | Summary of Liver Function Tests: Normal Range Categories by Country | This will include central lab data only | SAC |
| 3.72. | Safety | LB17 | Summary of Change from Baseline in Liver Function Tests of Potential Clinical Importance by Country | ALP, ALT, AST, GGT, Total bilirubin and albumin concentrations and PT/INR. This will include central lab data only | SAC |
| 3.73. | Safety | LB17 | Summary of Change from Baseline in Laboratory Data of Potential Clinical Importance by Country | This will include all lab data with a change from baseline PCI flag – albumin, creatinine, haematocrit, haemoglobin, | SAC |
| **Fasting Lipid Profile** | | | | | |
| 3.74. | Safety | LB1 | Summary of Change from Baseline in Fasting Lipid Profile by Country | Include absolute values for both V2 and V3 (Baseline) as well as change from baseline for all other visits<br>Fasting Lipid Profile include:<br>LDL Cholesterol, HDL Cholesterol, Total Cholesterol, Triglycerides | SAC |
| **Vitamin Absorption** | | | | | |
| 3.75. | Safety | LB1 | Summary of Change from Baseline in Vitamin Absorption by Country | Include absolute values for both V2 and V3 (Baseline) as well as change from baseline for all other visits<br>Vitamin Absorption includes:<br>Vitamin A, D, E and K | SAC |

| Safety : Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Haematology** | | | | | |
| 3.76. | Safety | LB1 | Summary of Change from Baseline in Haematology Data by Country | Include absolute values for both V2 and V3 (Baseline) as well as change from baseline for all other visits | SAC |
| 3.77. | Safety | LB17 | Summary of Haematology Data of Potential Clinical Importance by Country | | SAC |
| **Urinalysis** | | | | | |
| 3.78. | Safety | UR1 | Summary of Urinalysis Results by Visit by Country | | SAC |
| **Vital Signs and ECGs** | | | | | |
| 3.79. | Safety | VS1 | Summary of Change from Baseline in Vital Signs by Visit by Country | Include absolute values for V1, V2 and V3 (Baseline) as well as change from baseline for all other visits. | SAC |

| Safety : Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 3.80. | Safety | EG1 | Summary of ECG Findings by Country | Summary (freq and %) of Participants with normal and abnormal ECG results at each visit | SAC |
| 3.81. | Safety | EG2 | Summary of Change from Baseline in ECG Values by Visit by Country | Include absolute values for V1, V2 and V3 (Baseline) as well as change from baseline for all other visits. | SAC |
| **GSRS** | | | | | |
| 3.82. | Safety | Eff_T1 | Summary of GSRS Scores by Country | As per Eff_T1, paginated by domain.<br><br>For V3 to V7 (inclusive) | SAC |
| 3.83. | Safety | Eff_T1 | Summary of Change from Baseline in GSRS Scores by Country | As per Eff_T1, paginated by domain.<br><br>For V4 to V7 (inclusive) | SAC |
| 3.84. | Safety | EFF_T16 | Summary of Weekly e-diary GI symptoms questions (categorical) by Country | The number and proportion of people choosing each of the categories for each of the symptoms questions.<br><br>As per Eff_T9 with a sub-table per question. | SAC |
| 3.85. | Safety | Eff_T1 | Summary of Weekly e-diary GI symptoms questions (0-10 NRS) by Country | | SAC |
| 3.86. | Safety | Eff_T1 | Summary of Change from Baseline in Weekly e-diary GI symptoms questions (0-10 NRS) by Country | | SAC |
| **Country** | | | | | |
| 3.87. | Safety | AE1 | Summary of All Adverse Events by Country | Number of adverse events by SOC and PT | SAC |

## 12.12.8. Safety Figures

| Safety : Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Laboratory Data** | | | | | |
| 3.1. | Safety | EG9 | Line Plot of Change from Baseline in Chemistry Data against Visit | | SAC |
| 3.2. | Safety | EG9 | Line Plot of Change from Baseline in Liver Function Test Data against Visit | | SAC |
| 3.3. | Safety | EG9 | Line Plot of Change from Baseline in Liver Function Test Data against Visit by Risk of Disease Progression | | SAC |
| 3.4. | Safety | EG9 | Line Plot of Change from Baseline in Fasting Lipid Profile against Visit | | SAC |
| 3.5. | Safety | EG9 | Line Plot of Change from Baseline in Vitamin Absorption against Visit | | SAC |
| 3.6. | Safety | EG9 | Line Plot of Change from Baseline in Haematology against Visit | | SAC |
| **ECGs and Vital Signs** | | | | | |
| 3.7. | Safety | EG9 | Line Plot of Change from Baseline in QTc against Visit. | | SAC |
| 3.8. | Safety | EG9 | Line Plot of Change from Baseline in Vital Signs against Visit | One plot for each vital signs element | SAC |
| **GSRS** | | | | | |
| 3.9. | Safety | Eff_F16 | Boxplot of GSRS Scores against Visit: Panelled by Domain | Category = visit<br>Group = trtcd<br>Use a different colour for each treatment group.<br>Programming notes: plot each domain on a page with the appropriate scale (these are not all on the same scale as the range for the domains differs). | SAC |
| 3.10. | Safety | Eff_F7a | Mean Difference from Placebo in Change from Baseline in GSRS Score at Week 16 (V6), with 95% Confidence Intervals | Panelled by GSRS domain | SAC |
| 3.11. | Safety | Eff_F19a | Mean Difference from Placebo in Change from Baseline in GSRS Score at Week 16 (V6), with 95% Confidence Intervals by Country | Panelled by GSRS domain<br>Programming notes: Change symbols to different shapes | SAC |

| Safety : Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Special Interest AEs Diarrhoea** | | | | | |
| 3.12. | Safety | TTE10 | Kaplan-Meier Plot of Time to Diarrhoea | y-axis title should be "Probability without Diarrhoea" | SAC |
| 3.13. | Safety | Eff_F21 | Swimmer Plot of start of Diarrhoea and End of Diarrhoea by Treatment Group | Refer to mock ups. Each dose has one swimmer plot. Each lane is for each individual participant. Legend will be AE in question (during the ISP, MSP, and final study period). x-axis should be visit in weeks | HLR |
| 3.14. | Safety | TTE10 | Kaplan-Meier Plot of Time to Abdominal Pain | y-axis title to be "Probability without abdominal pain" | SAC |
| 3.15. | Safety | Eff_F21 | Swimmer Plot of start of Abdominal Pain and End of Abdominal Pain by Treatment Group | Refer to mock ups. Each dose has one swimmer plot. Each lane is for each individual participant. Legend will be AE in question (during the ISP, MSP, and final study period). x-axis should be visit in weeks. | SAC |

## 12.12.9. Pharmacokinetic Tables

| **Pharmacokinetic : Tables** | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| | | | | | |
| 4.1. | PK | PKCT1 | Summary of Linerixibat (GSK2330672) Plasma Pharmacokinetic Concentration-Actual Time Data (pg/mL) | Include all samples where linerixibat was measured (V3, V4, V5, and V6).<br>Record NQ and nearest Visit<br>See data specification section<br>By treatment group | SAC |
| 4.2. | PK | PKCT1 | Summary of UDCA Plasma Pharmacokinetic Concentration-Actual Time Data by (UDCA dose is same for all participants) | Include all samples for which UDCA was measured (V3, V4, V5 and V6).<br>Record NQ and nearest Visit<br>Need the unites for UDCA for PK<br>By treatment group | SAC |
| 4.3. | PK | PKCT1 | Summary of GUDCA Plasma Pharmacokinetic Concentration-Actual Time Data by (UDCA dose is same for all participants) | Include all samples for which GUDCA was measured (V3, V4, V5, and V6.<br>Record NQ and nearest Visit<br>Need the unites for GUDCA for PK<br>By treatment group | SAC |
| 4.4. | PK | PKCT1 | Summary of TUDCA Plasma Pharmacokinetic Concentration-Actual Time Data by (UDCA dose is same for all participants) | Include all samples for which TUDCA was measured (V3, V4, V5, and V6.<br>Record NQ and nearest Visit<br>Need the unites for TUDCA for PK<br>By treatment group | SAC |

| Pharmacokinetic : Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| | | | | | |
| 4.5. | PK | PKCT1 | Summary of Linerixibat (GSK2330672) Plasma Pharmacokinetic Concentration-Actual Time Data (pg/mL) by Country | Include all samples where linerixibat was measured (V3, V4, V5, and V6).<br>Record NQ and nearest Visit<br>See data specification section<br>By treatment group | SAC |
| 4.6. | PK | PKCT1 | Summary of UDCA Plasma Pharmacokinetic Concentration-Actual Time Data by (UDCA dose is same for all participants) by Country | Include all samples for which UDCA was measured (V3, V4, V5, and V6.<br>Record NQ and nearest Visit<br>By treatment group | SAC |
| 4.7. | PK | PKCT1 | Summary of GUDCA Plasma Pharmacokinetic Concentration-Actual Time Data by (UDCA dose is same for all participants) by Country | Include all samples for which GUDCA was measured (V3, V4, V5, and V6.<br>Record NQ and nearest Visit<br>By treatment group | SAC |
| 4.8. | PK | PKCT1 | Summary of TUDCA Plasma Pharmacokinetic Concentration-Actual Time Data by (UDCA dose is same for all participants) by Country | Include all samples for which TUDCA was measured (V3, V4, V5, and V6.<br>Record NQ and nearest Visit<br>By treatment group | SAC |
| 4.9. | PK | pkpt1 and pkpt3 | Summary of Derived PK Concentrations for Linerixibat | Separate linerixibat and linerixibat +UDCA cotreatment groups<br>(programming notes that is applied to all pkpt1 and pkpt3: pkpt1 is linear (un transformed data) and pkpt3 log transformed)<br>By treatment group | SAC |
| 4.10. | PK | pkpt1 and pkpt3 | Summary of Derived PK Concentrations for UDCA | Separate UDCA and UDCA +linerixibat cotreatment groups | SAC |

| **Pharmacokinetic : Tables** | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| | | | | | |
| 4.11. | PK | pkpt1 and pkpt3 | Summary of Derived PK Concentrations for GUDCA | Separate UDCA and UDCA +linerixibat cotreatment groups | SAC |
| 4.12. | PK | pkpt1 and pkpt3 | Summary of Derived PK Concentrations for TUDCA | Separate UDCA and UDCA +linerixibat cotreatment groups | SAC |

## 12.12.10. Pharmacokinetic Figures

| Pharmacokinetic : Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 4.1. | PK | PKCF1X | Line Plots of Individual Participant Linerixibat (GSK2330672) Plasma Pharmacokinetic Concentration-Actual Time Data (pg/mL) by Treatment Group | 1 plot per treatment group. Profile plot 1 line per participant<br>Header of plot should include in the legend a that there is a separate line for linerixibat alone and a separate line for linerixibat + UDCA.<br><br>Separate out Linerixibat Treatment only group from Linerixibat + UDCA co-treatment groups.<br>Linear and log plots | SAC |
| 4.2. | PK | PKCF1X | Line Plots of Individual Participant UDCA Plasma Pharmacokinetic Concentration-Actual Time Data by Treatment Group (UDCA dose is same for all participants) | 1 plot per treatment group. Profile plot 1 line per participant<br>Header of plot should include in the legend a that there is a separate line for linerixibat alone and a separate line for linerixibat + UDCA.<br><br>Separate out Linerixibat Treatment only group from Linerixibat + UDCA co-treatment groups<br>Linear and log plots | SAC |

| Pharmacokinetic : Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 4.3. | PK | PKCF1X | Line Plots of Individual Participant GUDCA Plasma Pharmacokinetic Concentration-Actual Time Data by Treatment Group (UDCA dose is same for all participants) | 1 plot per treatment group. Profile plot 1 line per participant<br>Header of plot should include in the legend a that there is a separate line for linerixibat alone and a separate line for linerixibat + UDCA.<br><br>Separate out Linerixibat Treatment only group from Linerixibat + UDCA co-treatment groups<br>Linear and log plots | SAC |
| 4.4. | PK | PKCF1X | Line Plots of Individual Participant TUDCA Plasma Pharmacokinetic Concentration-Actual Time Data by Treatment Group (UDCA dose is same for all participants) | 1 plot per treatment group. Profile plot 1 line per participant<br>Header of plot should include in the legend a that there is a separate line for linerixibat alone and a separate line for linerixibat + UDCA.<br><br>Separate out Linerixibat Treatment only group from Linerixibat + UDCA co-treatment groups<br>Linear and log plots | SAC |
| 4.5. | PK | PKCF1X | Line Plots of Individual Participant Linerixibat (GSK2330672) Plasma Pharmacokinetic Concentratin-Actual Time Data (pg/mL) by Treatment Group and by Country | 1 plot per treatment group. Profile plot 1 line per participant<br>Header of plot should include in the legend a that there is a separate line for linerixibat alone and a separate line for linerixibat + UDCA.<br><br>Separate out Linerixibat Treatment only group from Linerixibat + UDCA co-treatment groups.<br>Linear and log plots | SAC |

| Pharmacokinetic : Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 4.6. | PK | PKCF1X | Line Plots of Individual Participant UDCA Plasma Pharmacokinetic Concentration-Actual Time Data by Treatment Group (UDCA dose is same for all participants) and by Country | 1 plot per treatment group. Profile plot 1 line per participant<br>Header of plot should include in the legend a that there is a separate line for linerixibat alone and a separate line for linerixibat + UDCA.<br><br>Separate out Linerixibat Treatment only group from Linerixibat + UDCA co-treatment groups.<br>Linear and log plots | SAC |
| 4.7. | PK | PKCF1X | Line Plots of Individual Participant GUDCA Plasma Pharmacokinetic Concentration-Actual Time Data by Treatment Group (UDCA dose is same for all participants) and by Country | 1 plot per treatment group. Profile plot 1 line per participant<br>Header of plot should include in the legend a that there is a separate line for linerixibat alone and a separate line for linerixibat + UDCA.<br><br>Separate out Linerixibat Treatment only group from Linerixibat + UDCA co-treatment groups.<br>Linear and log plots | SAC |
| 4.8. | PK | PKCF1X | Line Plots of Individual Participant TUDCA Plasma Pharmacokinetic Concentration-Actual Time Data by Treatment Group (UDCA dose is same for all participants) and by Country | 1 plot per treatment group. Profile plot 1 line per participant<br>Header of plot should include in the legend a that there is a separate line for linerixibat alone and a separate line for linerixibat + UDCA.<br><br>Separate out Linerixibat Treatment only group from Linerixibat + UDCA co-treatment groups.<br>Linear and log plots | SAC |

| Pharmacokinetic : Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 4.9. | PK | Eff_F23 | Box Plots of Linerixibat Cave by Treatment Group and by Visit 4 and Visit 5 | X axis is treatment group<br>Y axis is C average<br>QD (low to high dose) then BID (low to high) doses left to right.<br>Legend is Visit 4 and Visit 5 | SAC |
| 4.10. | PK | Eff_F23 | Box Plots of UDCA Cave by Treatment Group and by Visit 4 and Visit 5 | X axis is treatment group<br>Y axis is C average<br>QD (low to high dose) then BID (low to high) doses left to right<br>DDI investigation<br>Legend is Visit 4 and Visit 5 | SAC |
| 4.11. | PK | Eff_F23 | Box Plots of GUDCA Cave by Treatment Group and by Visit 4 and Visit 5 | X axis is treatment group<br>Y axis is C average<br>QD (low to high dose) then BID (low to high) doses left to right<br>DDI investigation<br>Legend is Visit 4 and Visit 5 | SAC |
| 4.12. | PK | Eff_F23 | Box Plots of TUDCA Cave by Treatment Group and by Visit 4 and Visit 5 | X axis is treatment group<br>Y axis is C average<br>QD (low to high dose) then BID (low to high) doses left to right<br>DDI investigation<br>Legend is Visit 4 and Visit 5 | SAC |
| 4.13. | PK | Eff_F23 | Box Plots of Linerixibat Cave by Treatment Group by Visit 4 and Visit 5 and by Country | X axis is treatment group<br>Y axis is C average<br>QD (low to high dose) then BID (low to high) doses left to right<br>DDI investigation<br>Legend is Visit 4 and Visit 5 | SAC |

| Pharmacokinetic : Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 4.14. | PK | Eff_F23 | Box Plots of UDCA Cave by Treatment Group by Visit 4 and Visit 5 and by Country | X axis is treatment group<br>Y axis is C average<br>QD (low to high dose) then BID (low to high) doses left to right<br>DDI investigation<br>Legend is Visit 4 and Visit 5 | SAC |
| 4.15. | PK | Eff_F23 | Box Plots of GUDCA Cave by Treatment Group by Visit 4 and Visit 5 and by Country | X axis is treatment group<br>Y axis is C average<br>QD (low to high dose) then BID (low to high) doses left to right<br>DDI investigation<br>Legend is Visit 4 and Visit 5 | SAC |
| 4.16. | PK | Eff_F23 | Box Plots of TUDCA Cave by Treatment Group by Visit 4 and Visit 5 and by Country | X axis is treatment group<br>Y axis is C average<br>QD (low to high dose) then BID (low to high) doses left to right<br>DDI investigation<br>Legend is Visit 4 and Visit 5 | SAC |
| 4.17. | PK | PKCF1X | Line plot of change from baseline eGFR at V4 vs linerixibat Cave at V4.<br>(renal impairment exploration) | eGFR on y axis<br>C average on x axis<br>Participant must have a value for eGFR near V4 and linerixibat Cave near V4 to be included in plot<br>Linear and log plots | SAC |
| 4.18. | PK | PKCF1X | Line plot of change from baseline eGFR at V5 y-axis vs linerixibat Cave at V5.<br>(renal impairment exploration) | eGFR on y axis<br>C average on x axis<br>Participant must have a value for eGFR near V5 and linerixibat Cave near V5 to be included in plot<br>Linear and log plots | SAC |

## 12.12.11. Pharmacodynamic Biomarker Tables

| Pharmacodynamic Biomarker: Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Markers of Disease** | | | | | |
| **Biomarkers of PBC and Bile Acid Physiology** | | | | | |
| 5.1. | ITT | LB1 | Summary of change from baseline for PD biomarkers | Include absolute values for V3 (baseline), V4, V5, V6 and V7 as well as change from baseline and % change from baseline for all other visits.<br>PD biomarker data includes list in 10.1.1 and plasma lipid biomarkers. | SAC |
| 5.2. | ITT | LB1 | Summary of change from baseline for ALP for Subjects Taking UDCA | Include absolute values for V3 (baseline), V4, V5, V6 and V7 as well as change from baseline and % change from baseline for all other visits. | SAC |
| 5.3. | ITT | LB1 | Summary of change from baseline for PD biomarkers by Country | Include absolute values for V3 (baseline), V4, V5, V6 and V7 as well as change from baseline and % change from baseline for all other visits.<br>PD biomarker data includes list in 10.1.1 and plasma lipid biomarkers. | SAC |
| 5.4. | ITT | LB2 | Summary of change from V6 to V7 for PD biomarkers | Treat Visit 6 as baseline.<br>PD biomarker data includes list in 10.1.1 and plasma lipid biomarkers.<br>(off treatment washout observations) | SAC |
| 5.5. | ITT | LB2 | Summary of change from V6 to V7 for PD biomarkers by Country | Treat Visit 6 as baseline.<br>PD biomarker data includes list in 10.1.1 and plasma lipid biomarkers.<br>(off treatment washout observations) | SAC |

| Pharmacodynamic Biomarker: Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 5.6. | ITT | Eff_T2 | ANCOVA of mean change from baseline in TSBA at V6 (Week 16) | This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups. | SAC |
| 5.7. | ITT | Eff_T2 | ANCOVA of mean change from baseline in cholic acid (CA) at V6 (Week 16) | This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups. | SAC |
| 5.8. | ITT | Eff_T2 | ANCOVA of Mean Change from Baseline in Chenodeoycholic Acid (CDCA) at V6 (Week 16) | This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups. | SAC |
| 5.9. | ITT | Eff_T2 | ANCOVA of Mean Change from Baseline in Deoxycholic Acid (DCA) at V6 (Week 16) | This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups. | SAC |
| 5.10. | ITT | Eff_T2 | ANCOVA of Mean Change from Baseline in Serum C4 at V6 (Week 16) | This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups. | SAC |
| 5.11. | ITT | Eff_T2 | ANCOVA of Mean Change from Baseline in Autotaxin Concentrations at V6 (Week 16) | This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups. | SAC |
| 5.12. | ITT | Eff_T2 | ANCOVA of Mean Change from Baseline in FGF-19 at V6 (Week 16) | This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups. | SAC |

| Pharmacodynamic Biomarker: Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 5.13. | ITT | Eff_T2 | ANCOVA of Mean Change from Baseline in ELF Test at V6 (Week 16) | This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups. | SAC |
| 5.14. | ITT | Eff_T2 | ANCOVA of Mean Change from Baseline in Alkaline Phosphatase (ALP) at V6 (Week 16) | This will include the difference of all treatment groups to placebo and the difference between 90 mg BID and 180 mg QD dose groups. | SAC |
| 5.15. | ITT | Eff_T2 | ANCOVA of Mean Change from Baseline in Total Cholesterol at V6 (Week 16) | This will include the difference of all treatment groups to placebo | SAC |
| 5.16. | ITT | Eff_T2 | ANCOVA of Mean Change from Baseline in HDL Cholesterol at V6 (Week 16) | This will include the difference of all treatment groups to placebo | SAC |
| 5.17. | ITT | Eff_T2 | ANCOVA of Mean Change from Baseline in Direct LDL Cholesterol at V6 (Week 16) | This will include the difference of all treatment groups to placebo | SAC |
| 5.18. | ITT | Eff_T2 | ANCOVA of mean Mean Change from Baseline in Triglycerides at V6 (Week 16) | This will include the difference of all treatment groups to placebo | SAC |
| 5.19. | ITT | Eff_19 | Correlation Analysis of Baseline Biomarkers and Baseline Mean Worst Daily Itch Score by Treatment Group | List of PD biomarkers 10.1.1 and total bilirubin | SAC |
| 5.20. | ITT | Eff_19 | Correlation analysis of Mean Change from Baseline at V6 for PD Biomarkers and Mean Change from Baseline at V6 for Mean Worst Daily Itch Score by Treatment Group | List of PD biomarkers 10.1.1 and total bilirubin | SAC |
| 5.21. | ITT | Eff_19 | Correlation Analysis of change from baseline for ALP (V4-V3) and change from baseline for TSBA, CA, CDCA, DCA by Treatment Group for Visit 4 (Week 8) | V3 is baseline, | SAC |
| 5.22. | ITT | Eff_19 | Correlation Analysis of change from baseline for ALP (V5-V3) and change from baseline for TSBA, CA, CDCA, DCA by Treatment Group for Visit 5 (Week 12) | V3 is baseline, | SAC |

| Pharmacodynamic Biomarker: Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 5.23. | ITT | Eff_19 | Correlation Analysis of change from baseline for ALP(V6-V3) and change from baseline for TSBA, CA, CDCA, DCA by Treatment Group Visit 6 (Week 16) | V3 is baseline, | SAC |
| 5.24. | ITT | Eff_19 | Correlation Analysis change from baseline for ALP (V7-V6) and change from baseline for TSBA, CA, CDCA, DCA by Treatment Group for Visit 7 (Week 20) | V6 is baseline,<br>Off treatment washout view<br>From V6 to V7 | SAC |
| 5.25. | ITT | Eff_19 | Correlation analysis of change from baseline for ALP (V4-V3) and UDCA Cave, V4 concentration by Treatment Group (for placebo use concentrations zero for UDCA) for Visit 4 (Week 8) | V3 is baseline, | SAC |
| 5.26. | ITT | Eff_19 | Correlation Aanalysis of change from baseline for ALP (V5-V3) and UDCA Cave, V5 concentration by Treatment Group (for placebo use concentrations zero for UDCA) for Visit 5 (Week 12) | V3 is baseline, | SAC |
| 5.27. | ITT | Eff_19 | Correlation Analysis of change from baseline for ALP (V4-V3) and GUDCA Cave, V4 concentration by Treatment Group (for placebo use concentrations zero for GUDCA) for Visit 4 (Week 8) | V3 is baseline, | SAC |
| 5.28. | ITT | Eff_19 | Correlation Analysis of change from baseline for ALP (V5-V3) and GUDCA Cave, V5 concentration by Treatment Group (for placebo use concentrations zero for GUDCA) for Visit 5 (Week 12) | V3 is baseline, | SAC |
| 5.29. | ITT | Eff_19 | Correlation Analysis of change from baseline for ALP (V4-V3) and TUDCA Cave, V4 concentration by Treatment Group (for placebo use concentrations zero for TUDCA) for Visit 4 (Week 8) | V3 is baseline, | SAC |
| 5.30. | ITT | Eff_19 | Correlation Analysis of change from baseline for ALP (V5-V3) and TUDCA Cave, V5 concentration by Treatment Group (for placebo use concentration zero for TUDCA) for Visit 5 (Week 12) | V3 is baseline, | SAC |

## 12.12.12. Pharmacodynamic Biomarker Figures

| Pharmacodynamic Biomarker: Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Markers of Disease** | | | | | |
| **Biomarkers of PBC and Bile Acid Physiology** | | | | | |
| 5.1. | ITT | Eff_F4 | Line Plots of Mean PD Biomarkers Over Time by Treatment Group | Absolute values for V3, V4, V5 and V6. (mean +/- SD), PD biomarkers listed in 10.1.1<br><br>Programming notes: use the same shell but add +/- SD | SAC |
| 5.2. | ITT | Eff_F4 | Line Plots of Mean Change from Baseline for PD Biomarkers vs Visit by Treatment Group | V3 is baseline<br>(allows for view of PD response over time, PD biomarkers listed in 10.1.1)<br>Visit 4, Visit 5, and Visit 6 on one plot for each PD biomarker for Treatment Group (mean +/- SD)<br>(For example, PD CFB V4-V3 vs V4, V5-V3 vs V5, V6-V3 vs V6) | SAC |
| 5.3. | ITT | Eff_F24 | Bar Chart of mean change from baseline for PD biomarkers per Visit. | V3 is baseline<br>(dose response plots, PD biomarkers listed in 10.1.1)<br>One Bar chart with SD per visit for V4 V5, and V6; with treatment group on x-axis<br>Programming note: use the Eff_F24 shell and add SD<br>Treatment group colours should be consistent across all figures | SAC |

## 12.12.13. ICH Listings

| ICH : Listings | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 1. | Enrolled | ES7 | Listing of Reasons for Screen Failure | | SAC |
| 2. | Initial | ES7 | Listing of Reasons for Baseline Failure | | SAC |
| 3. | Initial | ES2 | Listing of Reasons for Premature Withdrawal from the study | Add a column with study period in which Participant was withdrawn (ISP, MSP, FSP) | SAC |
| 4. | Initial | SD2 | Listing of Reasons for Discontinuation of Study Treatment | Add a column with study period in which Participant was withdrawn (ISP, MSP, FSP) | SAC |
| 5. | Randomised | BL1 | Listing of Participants for Whom the Treatment Blind was Broken During the Study | | SAC |
| 6. | Randomised | TA1 | Listing of Planned and Actual Treatments | | SAC |
| 7. | Randomised | DV2 | Listing of Important Protocol Deviations | Add a column with study period in which protocol deviation occurred (ISP, MSP, FSP)<br>Please add a flag for ITT in this listing | SAC |
| 8. | Randomised | IE3 | Listing of Participants with Inclusion\Exclusion Criteria Deviations | | SAC |
| 9. | Enrolled | SP3 | Listing of Participants Excluded from any Population | | SAC |
| 10. | Safety | DM2 | Listing of Demographic Characteristics | | SAC |
| 11. | Safety | DM9 | Listing of Race | | SAC |
| 12. | Safety | CM3 | Listing of Concomitant Medications | | SAC |
| 13. | Safety | CM3 | Listing of Previous Treatments for PBC | | SAC |
| 14. | Safety | EX3 | Listing of Exposure | | SAC |
| 15. | Safety | EX3 | Listing of Compliance | Column for each of AM/PM and total compliance based on tablets and e-diary for each Visit in the Main Study Period (V3 to V6) | SAC |
| 16. | Safety | AE8 | Listing of all Adverse Events | | SAC |
| 17. | Safety | AE7 | Listing of Participant Numbers for Individual Adverse Events | | SAC |

| ICH : Listings | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| 18. | Safety | AE2 | Listing of Relationship Between Adverse Event System Organ Classes, Preferred Terms, and Verbatim Text | | SAC |
| 19. | Safety | AE9 | Listing of Serious Adverse events | | SAC |
| 20. | Safety | AE14 | Listing of Reasons for Considering as a Serious Adverse Event | | SAC |
| 21. | Safety | AE9 | Listing of Adverse Events Leading to Discontinuation of Study Drug or Withdrawal from the Study | | SAC |
| 22. | Safety | AE9 | Listing of Adverse Events of Special Interest | | SAC |
| 23. | Safety | AE9 | Listing of Other Significant Adverse Events | | SAC |
| 24. | Safety | MH2 | Listing of Medical Conditions for Participants with Liver Stopping Events | | SAC |
| 25. | Safety | SU2 | Listing of Substance Use for Participants with Liver Stopping Events | | SAC |
| 26. | Safety | LB6 | Listing of All Chemistry Data for Participants With at Least One Value of Potential Clinical Importance | | SAC |
| 27. | Safety | LB6 | Listing of All Haematology Data for Participants With at Least One Value of Potential Clinical Importance | | SAC |
| 28. | Safety | LB6 | Listing of All Urinalysis Data for Participants With at Least One Value of Potential Clinical Importance | | SAC |
| 29. | Safety | VS5 | Listing of All Vital Signs Data for Participants with at least one value of Potential Clinical Importance | | SAC |
| 30. | Safety | EG4 | Listing of All ECG Values for Participants With at Least One Value of Potential Clinical Importance | | SAC |
| 31. | Safety | EG6 | Listing of Abnormal ECG Findings | | SAC |

## 12.12.14. Non-ICH Listings

| Non-ICH : Listings | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **~~Analysis~~** | | | | | |
| 32. | ITT | | Listing of Risk Group Miss-stratification | List the risk group strata, ALP and total bilirubin values (both in terms of ratio of ULN) for Visit 2 and Visit 3 – for all those who were mis-stratified and/or changed risk group between V2 and V3 | SAC |
| **Safety** | | | | | |
| 33. | Safety | LB6 | Listing of Liver Chemistry at Screening and at each Visit during the study | | SAC |
| 34. | Safety | LB6 | Listing of All Faecal Occult Blood Data | | SAC |
| 35. | Safety | UR2a | Listing of All Urinalysis Data | | SAC |
| 36. | Safety | | Listing of GSRS Responses and Scores at each Week | Column for each domain | SAC |
| 37. | Safety | | Listing of Response to Symptoms Questions at each Week | Column for each question, categorical and NRS | SAC |
| 38. | Safety | | SAS output of Time to Diarrhoea | Include output of SAS statistical procedure | SAC |
| 39. | Safety | | SAS output of Time to Abdominal Pain | Include output of SAS statistical procedure | SAC |
| **eDiary Data** | | | | | |
| 40. | ITT | | Listing of Daily Worst Itch, Fatigue, Sleep and Overall Itch Scores | Column for each of AM, PM worst itch, AM, PM overall itch, , fatigue and sleep per Participant per day - for 7 days prior to each visit only | SAC |
| 41. | ITT | | Listing of mean worst daily itch, mean fatigue, mean sleep and mean overall itch scores per visit | Column for mean worst daily itch, mean sleep, mean fatigue and mean overall itch per Participant per visit | SAC |
| 42. | ITT | | SAS output for modelling of change from baseline in mean worst daily itch | Include output of SAS statistical procedure | SAC |
| 43. | PP | | SAS output for modelling of change from baseline in mean worst daily itch | Include output of SAS statistical procedure | SAC |

| Non-ICH : Listings | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Responders in Itch** | | | | | |
| 44. | ITT | | Listing of Participants meeting any of the Responder in Itch criteria at Week 16 (V6) and Number of Responder Days | A column for each responder criteria with Yes/No if met criteria at Week 16, a column for each responder definition with number of days criteria met between V3 and V6. | SAC |
| **Time to Worsening of Itch** | | | | | |
| 45. | ITT | | Listing of Time to Worsening of Itch | Column for time to worsening and indicator of whether censored or not | SAC |
| **Other eDiary itch questions** | | | | | |
| 46. | ITT | | Listing of Other e-diary questions (AM) | Column for response to each AM question in 7 days prior to each visit | SAC |
| 47. | ITT | | Listing of Other e-diary questions (PM) | Column for response to each AM question in 7 days prior to each visit | SAC |
| 48. | ITT | | Listing of Response to Follow-up telephone interview | | SAC |
| **Validated Questionnaires** | | | | | |
| **5-D Itch** | | | | | |
| 49. | ITT | | Listing of 5-D Itch Scores | Column for 5-D Itch domains and total score at each visit | SAC |
| **PBC-40** | | | | | |
| 50. | ITT | | Listing of PBC-40 Responses | Column for each question in the PBC-40, grouped by domain at each visit | SAC |
| **EQ-5D-5L** | | | | | |
| 51. | ITT | | Listing of EQ-5D-5L Responses | Column for each of the 5 questions in addition to a column for the VAS score and the derived Utility Index Score at each visit | SAC |
| **PGI** | | | | | |
| 52. | ITT | | Listing of PGI-C and PGI-S | Column for PGI-S and PGI-C responses at each visit | SAC |

| Non-ICH : Listings | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **BDI-II** | | | | | |
| 53. | ITT | | Listing of BDI-II Responses | Column for BDI-II score at each visit | SAC |
| **PTEA** | | | | | |
| 54. | ITT | | Listing of PTEA Responses | Column for each PTEA question | SAC |
| **Actigraphy** | | | | | |
| 55. | Actigraphy | | Listing of Actigraphy Data | Column for Rest Period Duration, WASO, Sleep Efficiency, Mean Duration of Scratching events per hour and Mean Number of scratching events per hour for each night there is data available. | SAC |
| **High Risk Population** | | | | | |
| 56. | High Risk | | Listing of Responders in Disease Progression Risk | Column for those in the high risk population – whether met disease progression risk responder criteria at each visit | SAC |
| 57. | High Risk | | Listing of Markers of Disease | Column for each marker of disease at each visit for all Participants in the high risk population | SAC |
| **Pharmacodynamic and Biomarkers of Disease** | | | | | |
| 58. | ITT | | Listing of PBC and Bile Acid Physiology | This will include:<br>Total serum bile acid, serum C4, serum autotaxin, FGF-19, ELF test, Serum Bile Acid Species | SAC |
| **PK** | | | | | |
| 59. | PK | | Listing of Plasma Pharmacokinetic Concentration-Time Data | Column for each of UDCA and GSK2330672 plasma concentration at each time point and visit. | SAC |
| 60. | PK | | Listing of PK Concentrations | Linerixibat, UDCA, GUDCA and TUDCA | SAC |
| **Medical Resource Utilization and Health Economics** | | | | | |
| 61. | Safety | Listing 32 | Listing of Medical Resource Utilization and Health Economics | This listing based on the hospital form | SAC |

## 12.13. Appendix 13: Example Mock Shells for Data Displays

Data Display Specification will be made available on request.

## 12.14. Appendix 14: Interim Analysis Results and update to the RAP

This appendix will be a separate document due to the sensitivity of interim analysis results. Please refer to Appendix 14 for additional details.

Signature Page for 201000 TMF-1771570 v1.0

| | |
|---|---|
| Reason for signing: Approved | Name: PPD<br>Role: Approver<br>Date of signature: 14-May-2020 15:12:01 GMT+0000 |

| | |
|---|---|
| Reason for signing: Approved | Name: PPD<br>Role: Approver<br>Date of signature: 14-May-2020 16:23:40 GMT+0000 |

Signature Page for TMF-1771570 v1.0

**CONFIDENTIAL**

**The GlaxoSmithKline group of companies** 201000

| | | |
|---|---|---|
| **Division** | : | Worldwide Development |
| **Information Type** | : | Reporting and Analysis Plan (RAP) Appendix 14 |

| | | |
|---|---|---|
| **Title** | : | Reporting and Analysis Plan for Reporting and Analysis Plan for Study 201000: A randomised, double-blind, multi-dose, placebo-controlled study to evaluate the efficacy, safety and tolerability of GSK2330672 administration for the treatment of pruritus in participants with primary biliary cholangitis. |
| **Compound Number** | : | GSK2330672 |
| **Effective Date** | : | 11-May-2020 |

**Description:**

- This RAP Appendix 14 contains sensitive information based on the Interim Analysis (IA) of the study. The Appendix will be reviewed and approved by the study team members who were unblinded to the aggregated IA results.

**RAP Author(s):**

| Approver | Date | Approval Method |
|---|---|---|
| PPD<br>Principal Statistician (Future Pipelines Discovery, Clinical Statistics) | 7-May-2020 | PPD |

**RAP Team Approvals:**

| Approver | Date | Approval Method |
|---|---|---|
| PPD<br>Medicine Development Leader | 8-May-2020 | email |
| PPD<br>Principal Programmer, Future Pipelines Discover Clinical Programming | 11-May-2020 | email |
| PPD<br>Director, Clinical Pharmacology Modeling and Simulation, CPMS | 8-May-2020 | email |

**Clinical Statistics and Clinical Programming Line Approvals:**

| Approver | Date | Approval Method |
|---|---|---|
| PPD<br>Director, Future Pipelines Discovery, Clinical Statistics | 8-May-2020 | email |

## 12.14. Appendix 14: Interim Analysis Results and update to RAP

This appendix contains an update to RAP (Reporting and Analysis Plan_Study201000) which is based on interim analysis (IA) results. This document reflects the IA decision based on the IA results; the outcome of instream decisions on future treatment groups, and this information is not revealed to the entire study team. To maintain concealment of the sensitive information that the interim analysis decision documents contain, this document will be stored in a restricted area outside the Pharma TMF until data base freeze.

The IA results are in "GSK2330672-201000 Interim Analysis (IA) Details decision July 2018" document.

- The pre-specified decision criteria are listed in the protocol under Section 10.3.4 and based on the IA results the team agreed to select decision criteria #3 and #4 and proceed with the trial accordingly.
- To better characterize the dose-response model for BID dosing and maximize information the following it was decided:
    - Drop 20 mg QD
    - Add 40 mg BID
    - Increase sample size to 140 subjects randomized
        - Ensure n = 20 for 40 mg BID
        - Increase n for 90 mg BID doses

As result, randomization schedule #7 (Drop 20mg QD and add 40mg BID, complete study with: 90mg QD, 180mg QD, 40mg BID and 90mg BID + placebo (increase N to 140)) was implemented.

This appendix will be reviewed and approved by the study team who were unblinded to the aggregated IA results.

All tables, figures and listings under the current RAP document will be updated for study population, safety, efficacy, PK, and biomarkers to include the new dose level 40 mg BID in addition to the 5 dose levels.

## 2.6 Considerations for Data analyses and Data Handling Conventions

### 5.1. Study Treatment & Sub-group Display Descriptors

| Treatment Group Descriptions | | | |
|---|---|---|---|
| [RandAll NG / FSO Randomization System] | | Data Displays for Reporting | |
| Code | Description | Description | Order in TLF |
| AA | Placebo for 16 weeks | Placebo | 1 |
| BB | GSK2330672 20 mg once daily for 12 weeks followed by placebo for 4 weeks | GSK2330672 20 mg QD | 2 |
| CC | GSK2330672 90 mg once daily for 12 weeks followed by placebo for 4 weeks | GSK2330672 90 mg QD | 3 |
| DD | GSK2330672 180 mg once daily for 12 weeks followed by placebo for 4 weeks | GSK2330672 180 mg QD | 4 |
| II | GSK2330672 40 mg twice daily for 12 weeks followed by placebo for 4 weeks | GSK2330672 40 mg BID | 5 |
| EE | GSK2330672 90 mg twice daily for 12 weeks followed by placebo for 4 weeks | GSK2330672 90 mg BID | 6 |

### 12.6.2 Study Population

| Treatment Compliance Number of Tablets |
|---|
| Treatment compliance calculations for study participant who were randomized using randomization schedule 2 up to IA should follow Section 12.6.2. under treatment compliance of the RAP. Randomization schedule 2 used 4 digits randomization number.<br><br>Post IA participant who were randomized using randomization schedule 7 (used 5 digits randomization number), treatment compliance will be calculated as following: Compliance will be calculated based on the number of tablets dispensed and returned within the MSP (dispensed at V3, V4 and V5 and returned at V4, V5 and V6).<br><br>**MSP Compliance** =100 * (total number of tablets dispensed from V3 to V6 – total number of tablets returned from V3 to V6) / (expected number of tablets to be taken) Where expected number of tablets to be taken = 8 tablets per day * (V6 –V3).<br><br>If a participant is lost to follow-up, compliance will be calculated up to last Visit attended, assuming no tablets were taken from any bottles dispensed at this visit. These participants should 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 MSP.

For example, if a participant does not return for V5 their compliance will be calculated as:

100 * (total number of tablets dispensed from V3 to V4 – total number of tablets returned from V3 to V4) / (expected number of tablets to be taken)
Where expected number of tablets to be taken = 8 tablets per day * (V4 –V3).

With a flag added to say that the participant dropped out after Visit 4.

However, compliance based on number of tablets will also be calculated for the containers marked as AM (2 containers, 4 tablets a day) and PM bottles (2 container, 4 tablets a day) over the duration of the main study period (Exposure start date at V3 to Exposure end date at V6).

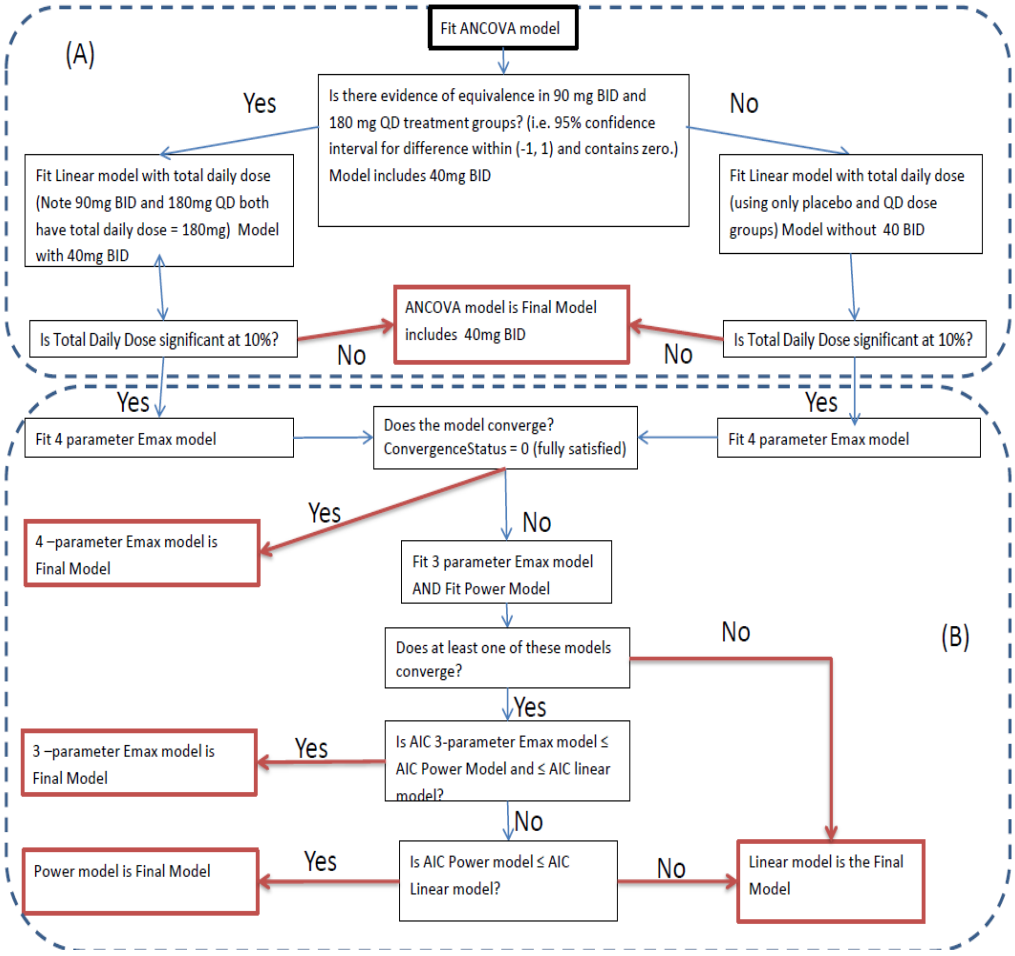
Compliance will be calculated similarly for the initial study period (V2 to V3) and final study period (V6 to V7).

#### 7.1.4.1. Statistical Methodology Specification

**Primary efficacy analysis**:

All primary efficacy analysis outputs will include the 40mg BID dose. Figure 1 under Section 7.1.4.1 is updated to include 40mg BID as below. Note the 4 parameter Emax model in part B of the figure will include 40mg BID.

**Figure 1 Flow chart of the model fitting process**




The primary efficacy analysis is updated as below (Please see example shells). In addition to the tables and figure that are listed under Section 12.12.5 and Section 12.12.6. in the RAP Table (2.6) and Figure (2.8) are added to the current primary efficacy outputs both outputs are to examine the BID doses only.

## Efficacy Tables

| Efficacy: Tables | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Mean Worst Daily Itch Score** | | | | | |
| 2.6 | ITT | Eff_T3a | ANCOVA Analysis of Change from Baseline in Mean Worst Daily Itch Score at Week 16 (V6).<br>Placebo and BID Doses | Mean difference for each total daily dose level from placebo. Note Include Placebo, 40mg BID, and 90 mg BID group. | SAC |

## Efficacy Figures

| Efficacy: Figures | | | | | |
|---|---|---|---|---|---|
| **No.** | **Population** | **IDSL / TST ID / Example Shell** | **Title** | **Programming Notes** | **Deliverable [Priority]** |
| **Mean Worst Daily Itch** | | | | | |
| 2.8. | ITT | Eff_F6 | ANCOVA Model: Estimated Mean Difference from Placebo in Change from Baseline in Mean Worst Daily Itch at Week 16 (V6) with 95% Confidence Intervals | Include ref line at y=0<br>Programming note: BID doses only,<br>refer to the shell in the RAP. Label X-axis "BID Treatment Group". | SAC |

**CONFIDENTIAL**

201000

Example : Eff_T2
Protocol : 201000
Population: [ITT]



Table X.X: Preliminary ANCOVA Analysis of Change from Baseline in Mean Worst Daily Itch at Week 16 (V6)

| | | | Change from Baseline | | Difference vs. Placebo | | Difference vs. GSK2330672 90 mg BID | |
|---|---|---|---|---|---|---|---|---|
| **Treatment Group** | **N** | **n** | **LS Mean** | **95% CI** | **LS Mean** | **95% CI [1]** | **LS Mean** | **95% CI [2]** |
| Placebo | X | X | x.x | (x.x, x.x) | NA | | NA | |
| 20 mg QD | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) | NA | |
| 90 mg QD | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) | NA | |
| 180 mg QD | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) | x.x | (x.x, x.x) |
| 40 mg BID | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) | x.x | (x.x, x.x) |
| 90 mg BID | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) | NA | |

[1] Mean difference < 0 favours active treatment, while mean difference > 0 favours placebo.

[2] If 95% CI is within (-1, 1) and includes zero, then the 90 mg BID and 180 mg QD can be thought of as equivalent (90 mg BID included in dose-response modelling at total daily dose of 180 mg).

Note: Mean Worst Daily Itch is the average of the worst daily scores provided in the 7 days prior to visit.

Note: The ANCOVA model includes Treatment Group and centred Mean Worst Daily Itch at Baseline.

**CONFIDENTIAL**

201000

Example : Eff_T3a



Protocol : 201000
Population : [ITT]

Table X.X: ANCOVA Analysis of Change from Baseline in Mean Worst Daily Itch at Week 16 (V6)
Placebo and BID Dose

| | | | Change from Baseline | | Difference vs. Placebo | |
|---|---|---|---|---|---|---|
| **Treatment Group** | **N** | **n** | **LS Mean** | **95% CI** | **LS Mean** | **95% CI[1]** |
| Placebo | X | X | x.x | (x.x, x.x) | NA | |
| 40 mg BID | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) |
| 90 mg BID | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) |

[1] Mean difference < 0 favours active treatment, while mean difference > 0 favours placebo.

Note: Mean Worst Daily Itch is the average of the worst daily scores provided in the 7 days prior to visit.

Note: The ANCOVA model includes Treatment Group and centred Mean Worst Daily Itch at Baseline.

N is the number in the population for each treatment group and n is the number of Participants with non-missing data available for analysis.

**CONFIDENTIAL**

201000

Example : Eff_T4c



Protocol : 201000
Population : [ITT]

Table X.X: Primary Analysis of Change from Baseline in Mean Worst Daily Itch at Week 16 (V6)

| | | | Change from Baseline | | Difference vs. Placebo [1] | | Probability Treatment Difference >= 2 | Mininally Effective Dose mg [2] |
|---|---|---|---|---|---|---|---|---|
| Treatment Group | N | n | LS Mean | 95% CI | LS Mean | 95% CI | P-value | |
| Placebo | X | X | x.x | (x.x, x.x) | NA | | | |
| 20 mg QD | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) | | |
| 90 mg QD | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) | | |
| 180 mg QD | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) | | |
| 40 mg BID | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) | | |
| 90 mg BID | X | X | x.x | (x.x, x.x) | x.x | (x.x, x.x) | | |

[1] Mean difference < 0 favours active treatment, while mean difference > 0 favours placebo.

[2] Minimally Effective Dose is the lowest dose treatment group at which the change from baseline in Mean Worst Daily Itch is at least 2.

Note: Mean Worst Daily Itch is the average of the worst daily scores provided in the 7 days prior to visit.

Note: The ANCOVA model includes Treatment Group and centred Mean Worst Daily Itch at Baseline.

*Programming Note: The final column 'probability treatment difference >=2' should only be included for the mean worst daily itch analysis. See RAP for details on how to calculate these columns.*

*Programming Note: This table will only be displayed if the ANCOVA is chosen as the best model (i.e. the dose-response modelling fails).*