PROTOCOL: MRX-502

TITLE:	MRX-502: Randomized Double-blind Placebo-controlled Phase 3 Study to Evaluate the Efficacy and Safety of **Maralixibat** in the Treatment of Subjects with Progressive Familial Intrahepatic **Ch**olestasis (**PFIC**) – **MARCH-PFIC**
SHORT TITLE:	A Placebo-controlled study of **Maralixibat** in Subjects with Progressive Familial Intrahepatic **Ch**olestasis (**MARCH-PFIC**)
DRUG:	Maralixibat (formerly SHP625 or LUM001)
IND:	119916
EudraCT:	2019-001211-22
SPONSOR:	Mirum Pharmaceuticals Inc. 950 Tower Lane Foster City, CA 94404
PRINCIPAL/ COORDINATING INVESTIGATOR:	Alexander Miethke, MD
PROTOCOL HISTORY:	Amendment 4, 10 May 2022 Amendment 3, 16 Jun 2020 Amendment 2, 22 Nov 2019 Amendment 1: 29 Apr 2019 Version 1: 14 Mar 2019

PROTOCOL SIGNATURE PAGE

Sponsor's (Mirum) Approval

Signature: DocuSigned by: Anamaria Lascau 7F5B41735F9E43A...	**Date:** 5/12/2022
Anamaria Lascau, MD Medical Lead	

Investigator's Acknowledgement

I have read this protocol for Mirum Pharmaceuticals Study MARCH-PFIC.

Title: MRX-502: Randomized double-blind placebo-controlled Phase 3 study to evaluate the efficacy and safety of maralixibat in the treatment of subjects with Progressive Familial Intrahepatic Cholestasis (PFIC) – MARCH-PFIC.

I have fully discussed the objective(s) of this study and the contents of this protocol with the sponsor's representative.

I understand that the information in this protocol is confidential and should not be disclosed, other than to those directly involved in the execution or the scientific/ethical review of the study, without written authorization from the sponsor. It is, however, permissible to provide the information contained herein to a subject in order to obtain their consent to participate.

I agree to conduct this study according to this protocol and to comply with its requirements, subject to ethical and safety considerations and guidelines, and to conduct the study in accordance with International Conference on Harmonisation guidelines on Good Clinical Practice and with the applicable regulatory requirements.

I understand that failure to comply with the requirements of the protocol may lead to the termination of my participation as an investigator for this study.

I understand that the sponsor may decide to suspend or prematurely terminate the study at any time for whatever reason; such a decision will be communicated to me in writing. Conversely, should I decide to withdraw from execution of the study I will communicate my intention immediately in writing to the sponsor.

Investigator Name and Address: (please hand print or type)	

Signature: _____ **Date:** _____

SUMMARY OF CHANGES FROM PREVIOUS VERSION

Changes are presented in ~~strikethrough~~ for deleted text and **bold** for added text:

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Synopsis</td><td>Text revised</td><td>Updated to reflect changes in the protocol</td></tr>
<tr><td>Section 2.1 Rationale for the Study</td><td>Updated text</td><td>Clarification</td></tr>
<tr><td colspan="3">~~There is currently~~ As of study initiation, there is no approved treatment for PFIC, and available medical approaches have limited efficacy.

The current study is designed to confirm the effect of higher doses of maralixibat ~~on pruritus, growth parameters~~ in PFIC2 and ~~relevant serum markers of the disease (ALT, TSB, and sBA) by using higher doses than those used to date in Study LUM001-501~~ assess it effect on a broader PFIC population. Because all treatment responders in Study LUM001-501 were subjects with PFIC2, the primary and secondary efficacy analyses will be limited to this subpopulation. ~~The current study therefore aims to confirm the efficacy and safety of maralixibat and to support a regulatory approval for the treatment of pruritus in PFIC 2 patients.~~

The ~~supplemental cohort will be used to assess the~~ efficacy and safety of maralixibat in PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) and other ~~PFIC~~ patient subpopulations will also be assessed based on the rationale that maralixibat also has the potential to lead to sBA and pruritus reduction in ~~those~~ the broader PFIC population and its subtypes.</td></tr>
<tr><td>Section 2.2.1 Primary Objective</td><td>Updated text</td><td>To clarify population</td></tr>
<tr><td colspan="3">• To evaluate the efficacy of maralixibat ~~vs.~~ versus placebo on the severity of pruritus in participants with PFIC2 ~~the primary cohort~~</td></tr>
<tr><td>Section 2.2.2 Secondary Objectives</td><td>Text added</td><td>Secondary objectives and endpoints updated to better capture clinically meaningful measures in the primary cohort and the broader PFIC population. The hierarchical testing was updated accordingly</td></tr>
<tr><td colspan="3">~~• To evaluate the efficacy of maralixibat vs. placebo on the frequency of pruritus in the primary cohort~~

• To evaluate the efficacy of maralixibat versus~~.~~ placebo on total sBA levels in participants with PFIC2 ~~the primary cohort~~

• To evaluate the efficacy of maralixibat versus placebo on the severity of pruritus in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)

• To evaluate the efficacy of maralixibat versus placebo on total sBA levels in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 2.2.2 Secondary Objectives (continued)</td><td>Text added</td><td>Secondary objectives and endpoints updated to better capture clinically meaningful measures in the primary cohort and the broader PFIC population. The hierarchical testing was updated accordingly</td></tr>
<tr><td colspan="3">• To evaluate the efficacy of maralixibat versus placebo on the proportion of responders for the ItchRO (Obs) pruritus score in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
• To evaluate the efficacy of maralixibat versus placebo on the sBA responder rate in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
• To evaluate the safety, tolerability, and ~~PK~~pharmacokinetics of maralixibat ~~vs.~~versus placebo in all participants who receive at least 1 dose of study medication (safety population) ~~PFIC subjects~~</td></tr>
<tr><td>Section 2.2.3 Exploratory Objectives</td><td>Text revised and added</td><td>Exploratory objectives and endpoints updated to better capture clinically meaningful measures in the primary cohort, the broader PFIC population, and other subgroups of interest.</td></tr>
<tr><td colspan="3">Additional endpoints may be explored to further characterize the efficacy and safety of maralixibat relative to placebo.
• ~~To evaluate the efficacy of maralixibat vs placebo in the primary cohort on the following parameters:~~
o ~~Liver biochemistry~~
o ~~Other measures of pruritus~~
o ~~Health-related quality of life~~
o ~~Responder rates of various pruritus measures and biomarkers~~
o ~~Quality of sleep~~
o ~~Subject growth~~
o ~~Other Patient Reported Outcomes (PRO), including Caregiver Impression of Severity (CIS), Patient Impression of Severity of Pruritus (PIS), or ItchRO(Pt)~~
o ~~Markers of bile acid metabolism: bile acid subspecies, C4, (FGF-19), autotaxin~~
o ~~Noninvasive fibrosis markers: Fibrosis-4 (FIB-4), AST to Platelet Ratio Index (APRI)~~
o ~~Time to liver-associated events~~
o ~~To evaluate the impact of maralixibat on healthcare utilization~~
o ~~To evaluate the primary, secondary, and exploratory endpoints in the supplemental cohort~~</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 3.1 Study Design and Flow Chart</td><td>Text revised</td><td>For clarity</td></tr>
<tr><td colspan="3">The primary analysis will be conducted in the ITT population of subjects with documented biallelic mutations in ABCB11 (PFIC2), based on standard-of-care genotyping (primary cohort criteria).

Subjects with other PFIC subtypes (e.g., PFIC 1/3/4, or new PFIC mutation variants) or postsurgical subjects (e.g., after internal or external biliary diversion surgery or subjects with reversal of biliary diversion surgery) will be enrolled in a separate supplemental cohort and ~~analyzed~~ evaluated as part of secondary and ~~for~~ exploratory ~~purposes~~ analyses (see Section 9.8).</td></tr>
<tr><td>Section 6.2.2 Allocation of Treatment to Subjects</td><td>Text added</td><td>Text added to specify that the SAP will detail how the assignment of siblings occurs.</td></tr>
<tr><td colspan="3">Siblings will be randomized in a blinded manner to the same treatment group. Additional details will be included within the SAP.</td></tr>
<tr><td>Section 7.2.1 Genetic Testing Results</td><td>Text added</td><td>For clarity of PFIC subtypes</td></tr>
<tr><td colspan="3">ATP8B1 (PFIC1), ABCB11 (PFIC2), ABCB4 (PFIC3), ~~and~~ TJP2 (PFIC4), NR1H4 (PFIC5), and MYO5B (PFIC6) mutations are predictive of PFIC.</td></tr>
<tr><td>Section 9.5 Validation Analysis Plan</td><td>Added section</td><td>To include a blinded analysis to confirm the responder definition</td></tr>
<tr><td colspan="3">A blinded analysis of the ItchRO(Obs) instrument will be performed before database lock according to the validation analysis plan (VAP). The blinded analysis will be used to estimate a threshold of clinically meaningful change (i.e., responder definition) in the 4-week average morning ItchRO(Obs) pruritus scores. The analysis will be performed by a group independent from both the study team and Mirum Pharmaceuticals. Blinded data will be used for this analysis; data will be collapsed across treatment groups.
Details of the analysis will be included in a standalone VAP, which will describe analyses that will be used to assess the measurement properties of the ItchRO(Obs) in alignment with scientific recommendations outlined in the FDA's Patient Focused Drug Development Guidance 3 (i.e., reliability, validity, ability to detect change, and anchor-based meaningful within-person clinically meaningful change criteria). Criteria for defining within-person clinically meaningful change for the ItchRO(Obs) will be detailed in the SAP.</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 9.7 Sample Size Calculation and Power Considerations</td><td>Added and deleted text</td><td>To clarify sample size calculation. To remove power calculation for endpoints other than the primary endpoint. To clarify primary cohort.</td></tr>
<tr><td colspan="3">The sample size was calculated ~~for~~ to ensure enrollment of an adequate number of PFIC2 participants who fulfill criteria for the primary cohort (primary endpoint)~~the primary cohort, defined as all PFIC2 subjects~~. The sample size is chosen to provide at least 80% power for the analysis of the primary endpoint ~~in the primary cohort~~.

Including a 10% dropout rate based on the previous PFIC study and rounding up to the next even number, approximately 30 participants with PFIC2 ~~subjects~~ (15 subjects in each treatment group) will be randomized ~~in the primary cohort of the study~~ for the primary analysis.

~~As for the secondary efficacy endpoint as measured by the mean change from baseline in the 4-week average morning ItchRO(Obs) frequency score from Week 15 to Week 26, with a between-treatment group LS mean difference of 0.676, the pooled SD of 0.580, and the effect size of 1.166, a total of 26 complete subjects (13 subjects in each treatment group) will provide 81% power for the comparison between the maralixibat treatment group and placebo, based on a 2-sided, 2-sample t-test at the 0.05 level of significance. Similar assumptions were made as for the primary endpoint.~~</td></tr>
<tr><td>Section 9.8 Analysis Populations</td><td>Removed and added text</td><td>Removed text to clarify that the ITT Population will include all randomized subjects. Added text to clarify the cohorts for the primary analysis of the primary and secondary efficacy endpoints.</td></tr>
<tr><td colspan="3">The intent-to-treat (ITT) population will consist of all randomized subjects. ~~Additional subgroups of the ITT population will be specified based on the presence/absence of post-baseline efficacy assessments for particular efficacy endpoints.~~

The per protocol population will consist of all subjects in the ITT population who receive at least 1 dose of study medication and do not have any major protocol violations or deviations. Major protocol violations/deviations will be identified prior to database lock.

The primary and secondary efficacy endpoints will be analyzed for each of the analysis populations described above. The primary analysis on the primary ~~and secondary~~ efficacy endpoint~~s~~ will be conducted in participants with PFIC2 in the ITT population. The primary analysis on the secondary efficacy endpoints will be conducted in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) in the ITT population.

~~The~~ Any exploratory efficacy endpoints will be analyzed in the ITT population. Exploratory analyses may be performed on participants with PFIC2, PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6), and other PFIC populations of interest. The final list of exploratory endpoints will be described in the SAP.

~~Subjects in the ITT population with no post-baseline efficacy assessments for a particular efficacy endpoint will be excluded from the analyses of that efficacy endpoint.~~</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 9.8.1 PFIC Populations for Analyses</td><td>Section added</td><td>Added to provide definitions of analyses subgroups</td></tr>
<tr><td colspan="3">The following definitions will be used for the PFIC populations in the analyses of the various endpoints in this study:
Primary and Secondary Endpoints:
• PFIC2: Participants with PFIC2 who fulfill criteria for the primary cohort (see Section 4)
• PFIC: Participants with PFIC1, PFIC2 (who fulfill criteria for the primary cohort), PFIC3, PFIC4, PFIC5, and PFIC6.
Exploratory Endpoints:
PFIC2 and PFIC populations (as above) and
• PFIC1: Participants with PFIC1
• PFIC3: Participants with PFIC3
• PFIC4: Participants with PFIC4
• PFIC5: Participants with PFIC5
• PFIC6: Participants with PFIC6
• Truncated PFIC2: Participants with PFIC2 with mutations leading to premature protein truncation or failure of protein production
• All participants: All participants enrolled
Each of the above-defined PFIC populations, with the exception of the All participants, will exclude any participants with biliary surgery at baseline and genotype variations affecting only 1 allele (heterozygous).
In the PFIC population (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6), participants with biallelic disease causing variation in ATP8B1 (PFIC1), ABCB11 (PFIC2), ABCB4 (PFIC3), TJP2 (PFIC4), NR1H4 (PFIC5), and MYO5B (PFIC6) may be included. The number of participants with each variation will depend on the genotype prevalence within the enrolled study population.</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 9.10 Efficacy Analysis</td><td>Text deleted and added</td><td>Removed duplicative text where appropriate and clarified efficacy analyses to be performed.</td></tr>
<tr><td colspan="3">~~Efficacy analyses on the primary and secondary efficacy endpoints will be performed in the ITT and per protocol populations. Analyses will be conducted in the ITT population separately for the primary cohort, overall supplemental cohort, all cohorts combined, and in the PFIC1 and PFIC3 sub-cohort (separately) given sufficient sample size. Analysis of the primary and secondary efficacy endpoints will also be performed on the primary cohort in the per protocol population.~~

~~Analyses on the exploratory efficacy endpoints will be performed in the ITT population separately on the primary cohort, overall supplemental cohort, all cohorts combined, and in the PFIC1 and PFIC3 sub-cohorts (separately) given sufficient sample size.~~

The primary and secondary efficacy analysis will be performed for the ITT population. Analysis of the primary and secondary efficacy endpoints will also be performed in the per-protocol population. Analyses on the exploratory efficacy endpoints will be performed in the ITT population~~.separately on~~ Exploratory analyses may be performed in participants with PFIC2, PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6), and other PFIC populations of interest. Final list of exploratory endpoints will be described in the SAP.

~~Efficacy analyses will be conducted in the ITT population separately on the primary cohort, the PFIC cohort, PFIC1 sub-cohort, PFIC3 sub-cohort, and all cohorts combined. Analysis of the primary and secondary efficacy endpoints will also be performed on the primary cohort in the PP population~~.

~~Analyses on the exploratory efficacy endpoints will be performed in the ITT population separately on the primary cohort, the PFIC cohort, PFIC1 sub-cohort, PFIC3 sub-cohort, and all cohorts combined.~~

~~The primary efficacy analysis will be performed in the primary cohort for the ITT population.~~

For the PFIC population (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6), PFIC type will also be used as a covariate in the analysis. For this analysis, due to the small sample sizes, PFIC4, PFIC5, and PFIC6 will be grouped together.</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 9.10.1 Primary Efficacy Endpoint</td><td>Text deleted and text added</td><td>Updated to clarify the scoring of ItchRO(Obs) and ItchRO(Pt) severity scores, to anchor the timing of averages to the first dose of study drug, and to clarify how to construct monthly scores. Additionally, language regarding estimands and the handling of intercurrent events was included.</td></tr>
<tr><td colspan="3">The primary efficacy endpoint is defined as the mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26, using 4-week average morning ItchRO(Obs) severity scores.

Weekly scores will be calculated relative to the initiation of treatment. For each week, a weekly score will be calculated if at least 4 days' worth of data are provided. If more than 50% of scores are missing for a particular week, the weekly average score will be treated as missing. Average scores will be calculated for the following periods: Through Week 6 (42 days), Weeks 7-10, 11-14, 15-18, 19-22, and 23-26 (28 days). For each period, the average of the weekly scores will be calculated. At least 50% of the weekly scores must be available in order to compute an average for the period. ~~Post-baseline average morning ItchRO(Obs) severity scores will be calculated as the sum of the morning ItchRO(Obs) severity scores divided by the number of morning ItchRO(Obs) severity scores for the 6-week (42-day) time period during the dose-escalation phase (i.e., through Week 6) and the 4-week (28-day) time periods prior to each visit during the stable dosing phase (i.e., Weeks 7–10, 11–14, 15–18, 19–22, and 23–26).~~ The baseline average morning ItchRO(Obs) score is defined as the 4-week average morning ItchRO(Obs) severity score prior to the first dose of study medication.

~~If 25% or more morning ItchRO(Obs) severity scores for the 6- or 4-week treatment period before a given study visit are missing, the average morning ItchRO(Obs) severity score at that visit will be treated as missing.~~</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 9.10.1 Primary Efficacy Endpoint (continued)</td><td>Text deleted and text added</td><td>Updated to clarify the scoring of ItchRO(Obs) and ItchRO(Pt) severity scores, to anchor the timing of averages to the first dose of study drug, and to clarify how to construct monthly scores. Additionally, language regarding estimands and the handling of intercurrent events was included.</td></tr>
<tr><td colspan="3"><u>Primary Analysis for Primary Efficacy Endpoint</u>

For this study, the primary estimand is the improvement in pruritus measured as change from baseline in the average morning ItchRO(Obs) severity score in the maralixibat treatment group relative to the placebo group. In the course of the 26-week randomized treatment period, participants may be exposed to possible known or unknown inter-current events that could possibly impact the interpretation of the measures associated with the clinical question of interest, such as treatment discontinuation due to a specific adverse effect or perhaps a lack of effect. A treatment policy estimand will be utilized for the analysis where all data will be incorporated. Intercurrent events will be handled as follows:

• Change in or administration of any allowed concomitant medication will be handled using a treatment policy strategy.
• Administration of any prohibited concomitant medication will be handled using a hypothetical strategy assuming that the prohibited medications are not available and hence were not taken, so as to assess the treatment effect for the randomized treatment groups in addition to allowed concomitant medication.
• Discontinuation of randomized treatment due to all of the following reasons will be handled using a treatment policy strategy:
  ○ Lack of efficacy
  ○ Safety concerns
  ○ Lack of compliance to study treatment
• Discontinuation of randomized treatment due to other reasons (including loss to follow-up) will be handled using a hypothetical strategy, assuming that randomized treatment had not been discontinued.
• Liver transplant or death: Neither of these are expected in the study. In the unlikely event that either of these events occur, they will be analyzed using a composite strategy, assuming worse possible score for all subsequent assessments.

The primary analysis on the primary efficacy endpoint will be conducted ~~over~~ in participants with PFIC2 who fulfill criteria for primary cohort ~~the~~ in the ITT population.</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 9.10.2 Adjustment for Multiplicity</td><td>Added and deleted text</td><td>To prioritize clinically meaningful endpoints in the statistical testing hierarchy</td></tr>
<tr><td colspan="3">In order to maintain study-wide Type I error control, a hierarchical testing procedure will be used in the comparisons between maralixibat and placebo on the primary and secondary efficacy endpoints ~~on the primary cohort~~ in the ITT population. The hierarchical order for testing the null hypotheses is as follows:

~~(1) mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26, (2) mean change in the average morning ItchRO(Obs) frequency score between baseline and Week 15 through Week 26, (3) mean change in total sBA level between baseline and Week 26, and (4) the proportion of subjects who experience an sBA control (defined as a reduction to <102 µMol/L, or a reduction of >75%, or normalization) from baseline through Week 26.~~

1. Primary: Mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26 in participants with PFIC2
2. Secondary: Mean change in total sBA level between baseline and average of Weeks 18, 22, and 26 in PFIC2 participants who fulfill criteria for the primary cohort
3. Secondary: Mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26 in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
4. Secondary: Mean change in total sBA level between baseline and average of Weeks 18, 22, and 26 in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
5. Secondary: Proportion of ItchRO(Obs) responders from Week 15 to Week 26 in participants with PFIC2, using the average value from the three 4-week periods (Weeks 15-18, 19-22, and 23-26)
6. Secondary: Proportion of sBA responders from Week 18 to Week 26 in participants with PFIC2, using the average value from Weeks 18, 22, and 26 values
7. Secondary: Proportion of ItchRO(Obs) responders from Week 15 to Week 26 in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
8. Secondary: Proportion of sBA responders from Week 18 to Week 26 in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)

Because the testing is sequential, the type I error rate of α=0.05 is maintained. Failure at any stage in the sequence implies no type I error control for the additional subsequent tests. All p values for each comparison without adjustments will be provided in the summary tables for informational purposes.

The effect of such a procedure is that no confirmatory claims can be based on the endpoint(s) that have a rank lower than that endpoint whose null hypothesis was the first that could not be rejected. Those tests will be considered as exploratory and no formal conclusions will be drawn.

~~For Europe and the rest of the world excluding the United States, the hierarchical testing will be performed on all of the endpoints listed above, in the order listed. For the United States, sBA endpoints (endpoint numbers 2, 4, 8, and 10) will be excluded from the hierarchical testing.~~</td></tr>
</table>

Protocol Amendment		
Summary of Change(s) Since Last Version of Approved Protocol		
Amendment Number 4	**Amendment Date** **10 MAY 2022**	
Section(s) Affected by Change	Description of Change	Rationale
Section 9.10.3 Secondary Efficacy Endpoints	Added and deleted text	Secondary objectives and endpoints updated to better capture clinically meaningful measures in the primary cohort and the broader PFIC population. The hierarchical testing was updated accordingly.

The secondary efficacy endpoints are defined as:

- ~~Mean change in the average morning ItchRO(Obs) frequency score between baseline and Week 15 through Week 26, using 4-week average morning ItchRO(Obs) frequency scores.~~
- ~~Mean change in total sBA level between baseline and Week 26~~
- ~~Proportion of subjects who experience an sBA control (defined as a reduction to <102 µMol/L, or a reduction of >75%, or normalization) from baseline through Week 26 Mean change in total sBA level between baseline and average of Weeks 18, 22, and 26.~~
- Mean change in total sBA level between baseline and average of Weeks 18, 22, and 26
- **Mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26 in participants with PFIC**
- **Proportion of ItchRO(Obs) responders from Week 15 to Week 26**
  - **For the purpose of determining response, the average value from the three 4-week periods (Weeks 15-18, 19-22, and 23-26) will be computed and used.**
    <u>ItchRO(Obs) responder definition:</u>
    Criteria for clinically meaningful change in the ItchRO(Obs) (responder definition) will be established by the VAP and described in detail in the SAP.
- **Proportion of sBA responders from Week 18 to Week 26**
  - **For the purpose of determining response, the average value from Weeks 18, 22, and 26 values will be computed and used.**
    <u>sBA responder definition:</u>
    - **≥75% reduction in sBA**

    OR
    - **sBA level <102 µmol/L (applies only if baseline sBA level was >102 µmol/L)**

~~Post-baseline average morning ItchRO(Obs) frequency scores will be calculated as the sum of the morning ItchRO(Obs) frequency scores divided by the number of morning ItchRO(Obs) frequency scores for the 6-week (42-day) time period during the dose-escalation phase (i.e., through Week 6) and the 4-week (28-day) time periods prior to each visit during the stable dosing phase (i.e., Weeks 7–10, 11–14, 15–18, 19–22, and 23–26). The baseline average morning ItchRO(Obs) frequency score is defined as the 4-week average morning ItchRO(Obs) frequency score prior to the first dose of study medication.~~
~~If 25% or more morning ItchRO(Obs) frequency scores for the 6- or 4-week treatment period before a given study visit are missing, the average morning ItchRO(Obs) frequency score at that visit will be treated as missing.~~
Serum bile acid from each scheduled clinic visit will be used in the analysis of mean change from baseline in sBA.

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 9.10.3 Secondary Efficacy Endpoints</td><td>Added and deleted text</td><td>Updated the hypothesis tests to be performed due to the adjustment of the secondary endpoints.</td></tr>
<tr><td colspan="3"><u>Analysis for Secondary Efficacy Endpoints</u>

Secondary efficacy endpoints will be analyzed in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) as described in Section 9.10. Analysis similar to that described for the primary efficacy endpoint, including the sensitivity analyses, will be performed for each of the secondary efficacy endpoints. ~~Analyses will be conducted separately on the primary cohort, overall supplemental cohort, all cohorts combined, PFIC1 sub-cohort, and PFIC3 sub-cohort in the ITT population. Analysis of the secondary efficacy endpoints will also be performed on the primary cohort in the per protocol population. Analysis on the PFIC1 and PFIC3 sub-cohorts will only be performed given sufficient sample size.~~

For the binary (responder) outcome~~s of the proportion of subjects who experience an sBA control (defined as a reduction to <102 µMol/L, or a reduction of >75%, or normalization) from baseline through Week 26~~, the number and proportion of subjects that are considered a “responder” will be summarized by treatment group at the end of study visit (Week 26/EOT). Barnard’s exact test will be used to calculate the p-value for the difference between treatment groups.

The null hypotheses for the secondary efficacy endpoints of the equality of maralixibat and placebo are:
~~H_{02}: mean change in the average morning ItchRO(Obs) frequency score between baseline and Week 15 through Week 26 in the 2 treatment groups are equal~~
~~H_{03}: mean change in total sBA level between baseline and Week 26 in the 2 treatment groups are equal~~
~~H_{04}: proportion of subjects who experience an sBA control (defined as a reduction to <102 µMol/L, or a reduction of >75%, or normalization) from baseline through Week 26 in the 2 treatment groups are equal~~
H_{02}: mean change in total sBA level between baseline and average of Weeks 18, 22, and 26 on participants with PFIC in the 2 treatment groups are equal
H_{03}: mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26 on participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) in the 2 treatment groups are equal
H_{04}: mean change in total sBA level between baseline and average of Weeks 18, 22, and 26 on participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) in the 2 treatment groups are equal
H_{05}: proportion of ItchRO(Obs) responders from Week 15 to Week 26 on participants with PFIC2 in the 2 treatment groups are equal
H_{06}: proportion of sBA responders from Week 15 to Week 16 on participants with PFIC2 in the 2 treatment groups are equal, using the average from Weeks 18, 22, and 26 sBA values
H_{07}: proportion of ItchRO(Obs) responders from Week 15 to Week 26 on participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) in the 2 treatment groups are equal
H_{08}: proportion of sBA responders from Week 18 to Week 26 on participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) in the 2 treatment groups are equal</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 9.10.4 Exploratory Efficacy Endpoints</td><td>Text added and deleted</td><td>Exploratory endpoints updated to better capture clinically meaningful measures in the primary cohort, the broader PFIC population, and other subgroups of interest.</td></tr>
<tr><td colspan="3">~~The exploratory efficacy endpoints are defined as:~~ The final exploratory endpoint list will be described in detail in the SAP. Analyses on the exploratory efficacy endpoints will be performed in the ITT population. Exploratory analyses may be performed in participants with PFIC2, PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6), and other PFIC populations of interest as described in the SAP.
• Mean change in liver-associated laboratory test level between baseline and Week 15 through Week 26
• Mean change from baseline in liver-associated laboratory tests ~~(i.e., ALT, total and conjugated bilirubin)~~ to Weeks 2, 6, 10, 14, 18, 22, and 26
• Proportion of participants with elevated ~~baseline~~ liver biochemistry (i.e., >ULN at baseline), whose liver biochemistry normalize (≤ULN) at Weeks 2, 6, 10, 14, 18, 22, and 26
• Mean change from baseline in FIB-4, ~~and~~ APRI, MELD, and PELD score to Weeks 2, 6, 10, 14, 18, 22, and 26
• Proportion of days with a morning, evening, and highest daily ItchRO(Obs) severity score ≤1 at the participant level from baseline to Week 6 (overall and weekly), Weeks 7-10, 11-14, 15-18, 19-22, 23-26, and 15-26
• Proportion of days with a morning, evening, and highest daily ItchRO(Obs) frequency score ≤1 at the participant level from baseline to Week 6 (overall and weekly), Weeks 7-10, 11-14, 15-18, 19-22, 23-26, and 15-26
• Proportion of days with a morning, evening, and highest daily ItchRO(Pt) severity score ≤1 at the participant level from baseline to Week 6 (overall and weekly), Weeks 7-10, 11-14, 1518, 19-22, 23-26, and 15-26
• Proportion of days with a morning, evening, and highest daily ItchRO(Pt) frequency score ≤1 at the participant level from baseline to Week 6 (overall and weekly), Weeks 7-10, 11-14, 15-18, 19-22, 23-26, and 15-26
• Mean change in the average morning ItchRO(Obs) frequency score between baseline and Week 15 through Week 26
• Mean change from baseline in average (6-week average for Week 6 and 4-week average afterwards) morning, evening, and highest daily ItchRO(Obs) severity and frequency scores to Weeks 6, 10, 14, 18, 22, 26, and Weeks 15-26 combined</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 9.10.4 Exploratory Efficacy Endpoints (continued)</td><td>Text added and deleted</td><td>Exploratory endpoints updated to better capture clinically meaningful measures in the primary cohort, the broader PFIC population, and other subgroups of interest.</td></tr>
<tr><td colspan="3">

Mean change from baseline in weekly average morning, evening, and highest daily ItchRO(Obs) severity and frequency scores to each post-baseline week, and Weeks 15-26 combined
Mean change from baseline in average (6-week average for Week 6 and 4-week average afterwards) morning, evening, and highest daily ItchRO(Pt) severity and frequency scores to Weeks 6, 10, 14, 18, 22, 26, and Weeks 15-26 combined
Mean change from baseline in weekly average morning, evening, and highest daily ItchRO(Pt) severity and frequency scores to each post-baseline week, and Weeks 15-26 combined
Mean change from baseline in CSS, CIS, and PIS at Weeks 2, 4, 6, 10, 14, 18, 22, and 26
Proportion of participants in each change from baseline category of the CSS, CIS, and PIS at Weeks 2, 4, 6, 10, 14, 18, 22, and 26
Proportion of responders at Weeks 6, 10, 14, 18, 22, 26, and any of those weeks, where a responder is defined as having:

a) Clinically meaningful 4-week average morning ItchRO(Obs) severity decrease from baseline AND a normalization or reduction from baseline in sBA of ≥75%, or average sBA level <102 µmol/L
b) Clinically meaningful 4-week average morning ItchRO(Obs) severity decrease from baseline
c) a normalization or reduction from baseline in sBA of ≥75%, or average sBA level <102 µmol/L

Number and proportion of participants achieving morning ItchRO(Obs) severity scores ≤1 for more than 50% of the time from baseline to Week 6, Weeks 7-10, 11-14, 15-18, 1922, 23-26, and 15-26
Mean change over time in quality of life as measured by the PedsQL
Mean change from baseline in average sleep disturbance scores over time (4-week average score at Weeks 6, 10, 14, 18, 22, and 26), based on sleep-related questions on the EDQ(Obs) morning, PedsQL (parent), and PedsQL (child) questionnaires, individually

</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 9.10.4 Exploratory Efficacy Endpoints (continued)</td><td>Text added and deleted</td><td>Exploratory endpoints updated to better capture clinically meaningful measures in the primary cohort, the broader PFIC population, and other subgroups of interest.</td></tr>
<tr><td colspan="3">

Proportion of “no” answers for sleep-related questions on the ItchRO(Obs) morning, ItchRO(Obs) evening, and ItchRO(Pt) morning, individually, at the participant level for baseline to Week 6, Weeks 7-10, 11-14, 15-18, 19-22, 23-26, 15-26, and overall
Proportion of participants with improvement from baseline (i.e., change from baseline <0) in EDQ(Obs) sleep score at Weeks 6, 10, 14, 18, 22, and 26 (using 4-week average scores)
Proportion of EDQ sleep scores ≤2 at the participant level post-baseline through Week 26
Mean change from baseline in height z-score and weight z-score at Weeks 2, 4, 6, 10, 14, 18, 22, and 26
Proportion of participants who experience an improvement from baseline in height z-score or weight z-score (i.e., change from baseline >0) at Weeks 2, 4, 6, 10, 14, 18, 22, and 26
Mean change, and % change from baseline in total sBA level to Weeks 2, 6, 10, 14, 18, 22, and 26
Mean change from baseline in bile acid subspecies, C4, FGF-19, and autotaxin to Weeks 10, 18, and 26
Time to liver-associated events (PEBD surgery, listing for liver transplantation, liver decompensation [hepatic encephalopathy, variceal bleeding, ascites, and spontaneous bacterial peritonitis] events, HCC, death)
Correlation analyses between sBA and pruritus severity

The ItchRO(Obs) frequency scores will be calculated using the same approach as the ItchRO(Obs) severity scores.

The SAP will describe the subscales/domains, including the scoring algorithms, for each of the PedsQL questionnaires. Liver biochemistry (e.g., TSB, ALT), <s>and</s> sBA normalization levels, along with definitions of elevated liver biochemistry levels and liver disease severity scores, will be covered in the SAP.
</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Section 9.10.4 Exploratory Efficacy Endpoints</td><td>Added and deleted text</td><td>Simplified the language by removing duplicative descriptions of endpoints.</td></tr>
<tr><td colspan="3"><u>Analysis for Exploratory Efficacy Endpoints</u>

~~Analyses will be conducted~~ ~~separately on the primary cohort, overall supplemental cohort, all cohorts combined, PFIC1 sub-cohort, and PFIC3 sub-cohort in the ITT population. Analysis on the PFIC1 and PFIC3 sub-cohorts will only be performed given sufficient sample size.~~

~~Efficacy endpoints with binary outcomes include: 1) the proportion of responders, where a responder is defined as having a 4-week average morning ItchRO(Obs) severity decrease from baseline of ≥ 1.0 AND a normalization or reduction from baseline in sBA of ≥ 70%, 2) the proportion of subjects with elevated baseline liver biochemistry, whose liver biochemistry normalize, and 3) the proportion of subjects who experience an improvement from baseline in height z-score or weight z-score (i.e., change from baseline >0).~~ For ~~these~~ responder-type endpoints, the number and proportion of subjects that are considered a “responder” will be summarized by treatment group for each visit, including the end of study visit (Week 26/EOT). Barnard’s exact test will be used to calculate the p-value for the difference between treatment groups at each study visit.</td></tr>
<tr><td>Section 9.11 Safety and Tolerability Analysis</td><td>Updated text</td><td>To clarify subgroups analyzed</td></tr>
<tr><td colspan="3">Analyses will be conducted separately on ~~the~~ participants with PFIC2 ~~primary cohort~~ and ~~the~~ participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6). ~~All participants cohort.~~</td></tr>
<tr><td>Section 9.12.1 Healthcare Utilization</td><td>Added text</td><td>For consistency with synopsis</td></tr>
<tr><td colspan="3">Healthcare resource utilization between baseline and Week 26 will be analyzed.</td></tr>
</table>

<table>
<tr><th colspan="3">Protocol Amendment</th></tr>
<tr><th colspan="3">Summary of Change(s) Since Last Version of Approved Protocol</th></tr>
<tr><th>Amendment Number 4</th><th colspan="2">Amendment Date
10 MAY 2022</th></tr>
<tr><td>Section(s) Affected by Change</td><td>Description of Change</td><td>Rationale</td></tr>
<tr><td>Appendix 10 Management of Clinical Study Procedures and Participants during COVID-19 Pandemic or Other Force Majeure</td><td>Deleted wording regarding remote visit approval.
Updated SAE reporting language.
Added section for “Guidance Regarding Vaccination against COVID-19.”</td><td>To provide guidance regarding COVID-19 vaccinations.</td></tr>
<tr><td colspan="3"><s>For all remote visits, medical monitor and Mirum medical approval is required.</s>

SAEs are reportable to the Sponsor within sponsor immediately without undue delay after becoming aware of the event and no later than 24 hours of awareness.

Guidance Regarding Vaccination against COVID-19

• Participants should be requested to contact the site upon planned vaccination against COVID-19. The vaccine should be entered in the eCRF as a concomitant medication and any new or worsened symptom or sign recorded as an adverse event and assessed for seriousness. Maralixibat is taken orally, is minimally absorbed with systemic levels below the lower limit of quantification, and binds with and inhibits the ileal bile acid transporter (IBAT) in the distal ileum. The vaccine is administered and acts systemically; hence, no drug-drug interaction is expected, and no need for unblinding is foreseen.</td></tr>
<tr><td colspan="3">Minor formatting and editorial changes have been made throughout the protocol for clarity and consistency.</td></tr>
</table>

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

STUDY SYNOPSIS

<table>
<tr><td>Protocol Number: MRX-502</td><td>Drug: Maralixibat (formerly SHP625 or LUM001)</td></tr>
<tr><td colspan="2">Title of the Study: Randomized Double-blind Placebo-controlled Phase 3 Study to Evaluate the Efficacy and Safety of Maralixibat in the Treatment of Subjects with Progressive Familial Intrahepatic Cholestasis (PFIC) – MARCH-PFIC</td></tr>
<tr><td colspan="2">Short Title: A Placebo-controlled study of Maralixibat in Subjects with Progressive Familial Intrahepatic Cholestasis (PFIC)</td></tr>
<tr><td colspan="2">Number of Subjects (total and for each treatment arm):

Up to approximately 30 subjects with PFIC2 will be enrolled in the primary cohort to achieve 26 subjects (13 per treatment group) completing the study. Subjects in the primary cohort will be randomized 1:1 to the maralixibat or placebo treatment groups.

A supplemental cohort of up to 60 subjects with other PFIC subtypes will also be included and randomized 1:1 to the maralixibat or placebo treatment groups.</td></tr>
<tr><td colspan="2">Investigator(s): International, multicenter study</td></tr>
<tr><td colspan="2">Site(s) and Region(s):

Study will be conducted at multiple sites in North America, Europe, Asia, and South America. Other regions may be added, if necessary, for enrollment purposes.</td></tr>
<tr><td>Study Period (planned): 2019 to 2022</td><td>Clinical Phase: 3</td></tr>
<tr><td colspan="2">Objectives

Primary Objective
• To evaluate the efficacy of maralixibat versus placebo on the severity of pruritus in participants with PFIC2

Secondary Objectives
• To evaluate the efficacy of maralixibat versus placebo on total serum bile acid (sBA) levels in participants with PFIC2
• To evaluate the efficacy of maralixibat versus placebo on the severity of pruritus in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
• To evaluate the efficacy of maralixibat versus placebo on total sBA levels in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
• To evaluate the efficacy of maralixibat versus placebo on the proportion of responders for the ItchRO(Obs) pruritus score in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
• To evaluate the efficacy of maralixibat versus placebo on the sBA responder rate in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
• To evaluate the safety, tolerability, and pharmacokinetics of maralixibat versus placebo in all participants who receive at least 1 dose of study medication (safety population)

Exploratory Objectives

Additional endpoints may be explored to further characterize the efficacy and safety of maralixibat relative to placebo.</td></tr>
</table>

Rationale

As of study initiation, there is no treatment approved for PFIC, and available medical approaches have limited efficacy. Preliminary data from the first interim analyses from the Phase 2 Study LUM001-501 (an open-label study of the efficacy and long term safety of maralixibat in the treatment of cholestatic liver disease in pediatric subjects with PFIC) indicate that a subgroup of subjects with PFIC2 experienced clinically significant improvements following maralixibat treatment as manifested by a large reduction or normalization of sBA, reduction in pruritus (ItchRO[Obs]), and normalization of total serum bilirubin (TSB) and ALT for those with elevated baseline values. Also, twice-daily (BID) dosing led to a higher response rate compared with the previously used once-daily (QD) dosing. The current study is designed to confirm the effect of higher doses of maralixibat in PFIC2 and assess its effect on a broader PFIC population.

Investigational Product, Dose, and Mode of Administration

Subjects in the current study will receive maralixibat oral solution based on their individual body weight, up to 600 µg/kg BID. Fixed concentrations of maralixibat oral solutions (i.e., 5, 10, 15, and 20 mg/mL) will be used. To accurately measure the required volume, 0.5-, 1.0-, and 3.0-mL sized dosing dispensers will be provided. The investigational product (IP) requires refrigerated storage conditions (2°C–8°C).

As much as possible, the morning dose should be administered approximately 30 minutes before breakfast and the evening dose approximately 30 minutes before the main evening meal. The doses will be administered prior to meals rather than every 12 hours to better cover the luminal bile acid release associated with meals.

Methodology

This is a 6-month, international, multicenter, randomized, double-blind, placebo-controlled Phase 3 study in subjects with PFIC. The study will be followed by the long-term extension Study MRX-503, during which all subjects who complete Study MRX-502 will have the opportunity to be treated with maralixibat.

Subjects will be randomized in a 1:1 ratio to the maralixibat or placebo treatment groups, respectively, stratified by cohort. The cohorts are defined as follows:

- Primary cohort, defined as subjects with non-truncated PFIC2
- Supplemental cohort, defined as:
    - Subjects with PFIC1, PFIC3 or PFIC4
    - Subjects with PFIC after internal or external biliary diversion surgery or for whom internal or external biliary diversion surgery was reversed
    - Subjects with PFIC phenotype without a known mutation or with another known mutation not described above

  Within the supplemental cohort, subjects will be randomized 1:1 by the following subcohorts: a) PFIC1; b) PFIC3; c) all other supplemental cohort subjects.

The study will be divided into 4 parts:

1. Screening period (up to 6 weeks)
2. Dose Escalation period (4–6 weeks)
3. Stable Dosing period (20–22 weeks)
4. Safety Follow-up period (7 days; for subjects discontinuing early and for subjects not enrolling into the extension study MRX-503)

Subjects randomized to placebo will remain on placebo for the entire duration of the study. The Dose Escalation period will consist of the following weekly steps:

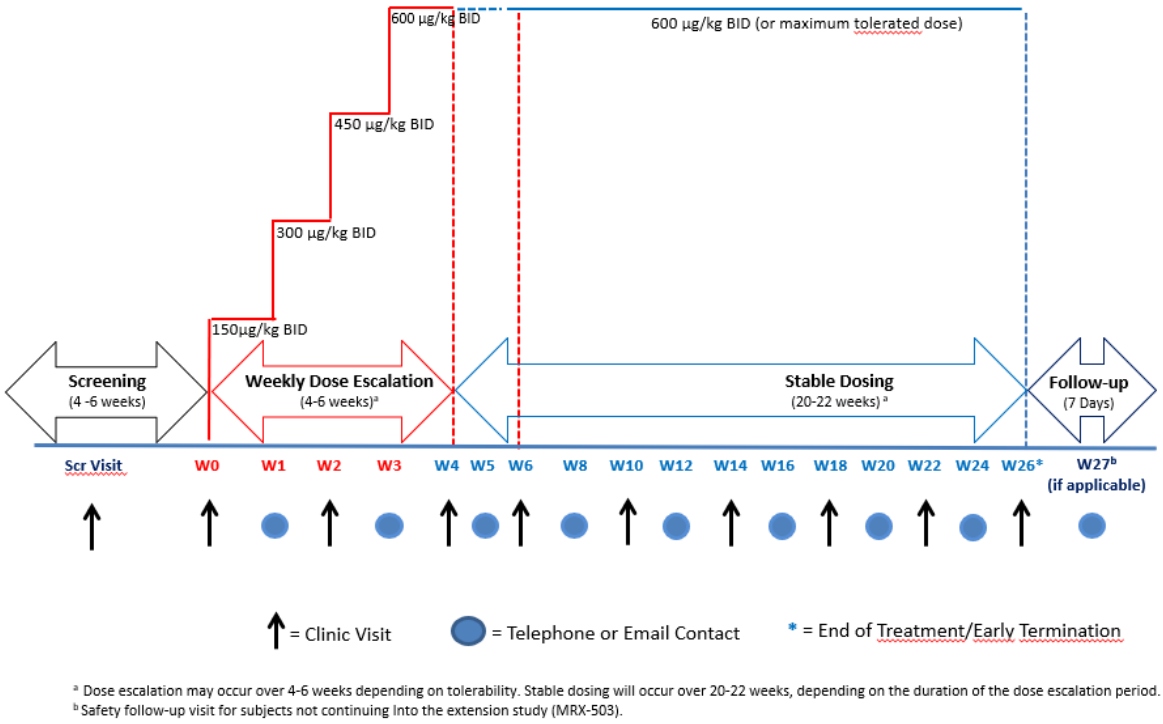
Dose level 1: 150 µg/kg maralixibat BID for 1 week

Dose level 2: 300 µg/kg maralixibat BID for 1week

Dose level 3: 450 µg/kg maralixibat BID for 1 week

Dose level 4: 600 µg/kg maralixibat BID for the remaining duration of the study

Figure 1 Study Design Flow Chart



[a] Dose escalation may occur over 4-6 weeks depending on tolerability. Stable dosing will occur over 20-22 weeks, depending on the duration of the dose escalation period.
[b] Safety follow-up visit for subjects not continuing into the extension study (MRX-503).

BID=twice daily; Scr=screening; W=week

Dose escalation should occur in the absence of major safety (e.g., liver parameters) or tolerability concerns (e.g., gastrointestinal [GI]-related treatment-emergent adverse events [TEAEs]) related or possibly related to study medication. Subjects with such safety concerns can be down-titrated to a lower, previously tolerated dose level for 1 week before continuing dose escalation. The minimum dose to continue in the study will be 150 µg/kg maralixibat BID; subjects who cannot tolerate this dose will be discontinued from the study. Investigators have until Week 6 to determine the maximum tolerated dose; if re-challenges or further dose escalations fail, the subject will remain on the maximum tolerated dose level for the remainder of the Stable Dosing period to complete up to 20 weeks of stable dosing. Dose reductions are allowed for the above-mentioned safety or tolerability reasons down to a minimum level of 150 µg/kg BID. Subjects who cannot tolerate this dose will be discontinued from the study.

During the Stable Dosing period, subjects will be allowed a maximum of 28 days (both morning and afternoon dose missed) of cumulative dose interruption. During the last 12 Weeks of the Stable Dosing period (Week 15-26), subjects will be allowed a maximum of 17 days (both morning and afternoon dose) of cumulative dose interruption, but not more than 7 consecutive days. Subjects may remain on study during dose interruptions and should continue to complete all regularly scheduled subject study visits and assessments. After any dose interruption, subjects will reinitiate dosing at their maximum tolerated dose. Any dose interruption longer than the maximum allowed will be documented as a protocol deviation.

Any subject who has over 56 days (both morning and afternoon dose) of cumulative dose interruption during the Stable Dosing period should be discontinued from the study.

Inclusion and Exclusion Criteria

Inclusion Criteria

1. Informed consent and assent (as applicable) per Institutional Review Board/Ethics Committee (IRB/EC)
2. Male or female subjects with a body weight ≥5 kg, who are ≥12 months and <18 years of age at time of baseline
3. Cholestasis as manifested by total sBA ≥3× ULN (applies to primary cohort only)
4. An average AM ItchRO(Obs) score ≥1.5 during 4 consecutive weeks of the screening period, leading to the baseline visit (Visit 1)
5. Completion of at least 21 valid* morning ItchRO(Obs) entries during 4 consecutive weeks of the screening period, leading to the baseline visit (Visit 1) (*valid=completed and not answered as "I don't know"; maximum allowed invalid reports=7, no more than 2 invalid reports during the last 7 days before randomization)
6. Diagnosis of PFIC based on the following:

   Chronic cholestasis as manifested by persistent (>6 months) pruritus in addition to biochemical abnormalities and/or pathological evidence of progressive liver disease

   and

   Primary Cohort:

   Subjects with genetic testing results consistent with biallelic disease-causing variation in *ABCB11* (PFIC2), based on standard-of-care genotyping

   Supplemental Cohort:

   i. Subjects with genetic testing results consistent with biallelic disease-causing variation in *ATP8B1* (PFIC1), *ABCB4* (PFIC3), or *TJP2* (PFIC4), based on standard-of-care genotyping

   ii. Subjects with PFIC phenotype without a known mutation or with another known mutation not described above or with intermittent cholestasis as manifested by fluctuating sBA levels

   iii. Subjects with PFIC after internal or external biliary diversion surgery or for whom internal or external biliary diversion surgery was reversed
7. Male and females of non-childbearing potential. Males and non-pregnant, non-lactating females of childbearing potential who are sexually active must agree to use acceptable methods of contraception during the study and 30 days following the last dose of the study medication. Females of childbearing potential must have a negative pregnancy test result.
8. Access to email or telephone for scheduled remote visits
9. Ability to read and understand the questionnaires (both caregivers and subjects above the age of assent)
10. Access to consistent caregiver(s) during the study
11. Subject and caregiver willingness to comply with all study visits and requirements

Exclusion Criteria

1. Predicted complete absence of bile salt excretion pump (BSEP) function based on the type of ABCB11 mutation (PFIC2), as determined by a standard-of-care genotyping (applies to primary cohort only). Subjects can enter the study in the supplemental cohort (under inclusion criteria 6.ii or 6.iii).

2. Recurrent intrahepatic cholestasis, indicated by a history of sBA levels <3x ULN or intermittent pruritus (applies to primary cohort only)
3. Current or recent history (<1 year) of atopic dermatitis or other non-cholestatic diseases associated with pruritus
4. History of surgical disruption of the enterohepatic circulation (applies to primary cohort only)
5. Chronic diarrhea requiring intravenous fluid or nutritional intervention for the diarrhea and/or its sequelae at screening or during the 6 months prior to screening
6. Previous or need for imminent liver transplant
7. Decompensated cirrhosis (international normalized ratio [INR] >1.5, albumin <30 g/L, history or presence of clinically significant ascites, variceal hemorrhage, and/or encephalopathy)
8. ALT or total serum bilirubin (TSB) >15× ULN at screening
9. Presence of other liver disease
10. Presence of any other disease or condition known to interfere with the absorption, distribution, metabolism or excretion of drugs, including bile salt metabolism in the intestine (e.g., inflammatory bowel disease), per investigator discretion
11. Possibly malignant liver mass on imaging, including screening ultrasound
12. Known diagnosis of human immunodeficiency virus (HIV) infection
13. Any prior cancer diagnosis (except for in situ carcinoma) within 5 years before the screening visit (Visit 0)
14. Any known history of alcohol or substance abuse
15. Administration of bile acids or lipid binding resins, or phenylbutyrates during the screening period
16. Criterion has been deleted as of Amendment 3.
17. Administration of any investigational drug, biologic, or medical device during the screening period
18. Previous use of an ileal bile acid transporter inhibitor (IBATi)
19. History of nonadherence to medical regimens, unreliability, medical condition, mental instability or cognitive impairment that, in the opinion of the investigator or sponsor medical monitor, could compromise the validity of informed consent, compromise the safety of the subject, or lead to nonadherence with the study protocol or inability to conduct the study procedures
20. Known hypersensitivity to maralixibat or any of its excipients

Enrollment in the study will be considered complete once the target number of subjects has been reached in the *primary cohort*. Enrollment into the supplemental cohort will be capped at a maximum of 60 subjects.

Maximum Duration of Subject Involvement in the Study:

- Planned duration of Screening period: up to 6 weeks
- Planned duration of Dose Escalation period: 4–6 weeks (maximum 6 weeks)
- Planned duration of Stable Dosing period: 20–22 weeks (minimum 20 weeks)
- Planned duration of Safety follow-up: 7 days (for subjects discontinuing early for reasons other than safety/tolerability and subjects not enrolling into the extension study MRX-503)

Active treatment duration: up to 26 weeks

Total study duration: up to 33 weeks

Endpoints and Statistical Analysis:

Subject Populations

- The **intent-to-treat (ITT) population** will consist of all randomized subjects.
- The **per-protocol population** will consist of all subjects in the ITT population who receive at least 1 dose of study medication and do not have any major protocol violations or deviations. Major protocol violations/deviations will be identified prior to database lock.
- The **safety population** will consist of all subjects who receive at least 1 dose of study medication.

Endpoints

Efficacy Endpoints

Primary Efficacy Endpoint

- Mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26, using 4-week average morning ItchRO(Obs) severity scores

Secondary Efficacy Endpoints

- Mean change in total sBA level between baseline and average of Weeks 18, 22, and 26
- Mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26 in participants with PFIC
- Proportion of ItchRO(Obs) responders from Week 15 to Week 26
    - For the purpose of determining response, the average value from the three 4-week periods (Weeks 15-18, 19-22, and 23-26) will be computed and used.
- Proportion of sBA responders from Week 18 to Week 26
    - For the purpose of determining response, the average value from Weeks 18, 22, and 26 values will be computed and used.

Safety & Tolerability Endpoints

- Treatment emergent AEs will be further summarized by severity and relationship to study medication. Adverse events related to study medication, AEs leading to withdrawal, SAEs, and deaths will be similarly summarized.
- Change from baseline in clinical laboratory values, physical examination findings (including body weight, height, and body mass index [BMI]), vital signs, and electrocardiogram (ECG) parameters

Other Endpoints

- Healthcare resource utilization between baseline and Week 26 will be analyzed.
- Systemic concentrations of maralixibat in plasma will be determined at pre-dose and at approximately 2.5 hours after the morning dose at Week 10 and Week 26 (EOT/ET)

The primary and secondary efficacy endpoints will be evaluated in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6).

All safety analyses will be performed in the safety population and separately in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, PFIC6).

In general, baseline is defined as all data collected at Visit 1 (Day 0) or the last assessment before the first dose of the study medication if the data are not collected at Visit 1 (Day 0). Baseline ItchRO and EDQ average scores are defined as the 4-week average score prior to the first dose of study medication.

Statistical Methodology for Primary Efficacy Endpoint

The primary efficacy endpoint is defined as the mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26, using 4-week average morning ItchRO(Obs) severity scores. The baseline average morning ItchRO(Obs) score is defined as the 4-week average morning ItchRO(Obs) severity score prior to the first dose of the study medication.

The primary analysis on the primary efficacy endpoint will be conducted in participants with PFIC2 who fulfill criteria for the primary cohort in the ITT population. A restricted maximum likelihood (REML)-based mixed-effects model for repeated measures (MMRM) will be used as the primary analysis method. The repeated measures include post-baseline visits during the dose escalation phase (i.e., Week 6) and stable dosing phase (i.e., Weeks 10, 14, 18, 22, and 26), with change from baseline in the 6- or 4-week average morning ItchRO(Obs) severity score as the dependent variable. The MMRM model will include the fixed, categorical effects of treatment group, visit, and treatment group-by-visit interaction as well as the continuous, fixed covariates of baseline 4-week average morning ItchRO(Obs) severity score and the baseline score-by-visit interaction. The unstructured variance/covariance matrix will be used to model the variances and covariances for the 6 time points (periods) included in the model. The unstructured variance/covariance does not impose any restrictions on the pattern of the matrix elements. Every attempt (e.g., relaxing the convergence criteria, increasing the iteration limit, choosing reasonable starting values for the estimates) will be made to ensure convergence using the unstructured modeling of within-subject correlations. The primary efficacy analysis will compare maralixibat and placebo using the contrast (difference in least squares [LS] mean) between treatment groups during the last 12 weeks of the study (i.e., Weeks 15 to 26 combined).

The null hypothesis of equal treatment effect will be rejected if the statistical analysis results in a 2-sided p-value for treatment over the last 12 weeks of the study is ≤ 0.05. Least squares means will be calculated for each treatment group for each post-baseline period in the model. The difference between maralixibat and placebo change from baseline in average morning ItchRO(Obs) severity score will be estimated, with the corresponding 2-sided 95% confidence interval (CI) constructed for each period, and the last 12 weeks (i.e., last 3 periods) of the study combined. The change from baseline LS means with standard error, 95% CI for the LS means, p-value for testing if the LS mean is zero, LS mean difference between treatment groups (maralixibat minus placebo) with standard error, 95% CI for the LS mean difference, and p-value for testing if the treatment LS means are equal will be presented.

Sensitivity and supportive analysis for the primary efficacy endpoint will be conducted in PFIC2 participants who fulfill criteria for the primary cohort in the ITT population using the MMRM model as the primary analysis.

The trial will be claimed successful if the hypothesis of no treatment effect on the primary efficacy endpoint in the ITT population is rejected at the 0.05 (2-sided) significance level.

Adjustment for Multiplicity

In order to maintain study-wide Type I error control, a hierarchical testing procedure will be used in the comparisons between maralixibat and placebo on the primary and secondary efficacy endpoints in the ITT population.

Analysis for Secondary Efficacy Endpoints

Analysis similar to that described for the primary efficacy endpoint, including the sensitivity analyses, will be performed for the secondary efficacy endpoints. Analyses will be conducted separately in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2. PFIC3. PFIC4, PFIC5, PFIC6) in the ITT population. Analysis of the secondary efficacy endpoints will also be performed on both populations in the per-protocol population.

For the binary (responder) outcomes, the number and proportion of subjects that are considered a "responder" will be summarized by treatment group at the end-of-study visit (Week 26/EOT). Barnard's exact test will be used to calculate the p-value for the difference between treatment groups.

All tests will be performed as 2-sided tests at the 0.05 level of significance.

Statistical Methodology for Exploratory Efficacy Endpoints

Any analyses on the exploratory efficacy endpoints will be performed in the ITT population. Exploratory analyses may be performed separately on participants with PFIC2, PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) and other PFIC populations of interest. All tests will be performed as 2-sided tests at the 0.05 level of significance.

Sensitivity analysis will not be performed on the exploratory endpoints.

Statistical Methodology for Safety Endpoints

All safety analyses will be performed in the safety population and separately in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, PFIC6). Safety measures will be summarized descriptively. For quantitative variables, descriptive statistics including number of observations, mean, median, standard deviation, minimum, and maximum will be presented for the observed and change from baseline (Visit 1) values at each study visit. Qualitative variables will be summarized using counts and percentages.

Adverse events and concomitant treatments will be coded with standard dictionaries and will be summarized by treatment received.

Statistical Methodology for Pharmacokinetic Endpoint(s)

Summary statistics (number of observations, mean, standard deviation, coefficient of variation, median, minimum, maximum, and geometric mean) will be determined for maralixibat concentrations at each nominal time point and sampling time.

ABBREVIATIONS

ABCB4	ATP binding cassette subfamily B member 4
ABCB11	ATP binding cassette subfamily B member 11
ADR	adverse drug reaction
AE	adverse event
AFP	α-fetoprotein
ALG	Alagille syndrome
ALP	alkaline phosphatase
ALT	alanine aminotransferase
APRI	AST-to-platelet ratio index
ASBTi	apical sodium-dependent bile acid transporter inhibitor
AST	aspartate aminotransferase
ATP	adenosine triphosphate
ATP8B1	ATPase phospholipid transporting 8B1
β-hCG	beta-human chorionic gonadotropin
BID	twice daily (dosing)
BL	baseline
BMI	body mass index
BSEP	Bile Salt Excretion Pump
C4	7α-hydroxy-4-cholesten-3-one
CBC	complete blood count
CI	confidence interval
CIC	Caregiver Impression of Change
CIS	Caregiver Impression of Severity
COA	Clinical Outcomes Assessment
CRA	clinical research associate
CRF	case report form
CRO	contract research organization
CSS	Clinician Scratch Scale
DMC	data monitoring committee
ECG	electrocardiogram
EDQ	exploratory diary questionnaire
EOT	end of treatment
ET	early termination
FDA	US Food and Drug Administration
FGF-19	fibroblast growth factor-19
GI	gastrointestinal
GCP	Good Clinical Practice
HCV	hepatitis C virus
HIV	human immunodeficiency virus
IB	Investigator's Brochure
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
INR	international normalized ratio
IP	investigational product
IRB	Institutional Review Board

ItchRO™	Itch Reported Outcome
ItchRO(Obs)™	Observer-reported Itch Reported Outcome Instrument
ItchRO(Pt)™	Patient-reported Itch Reported Outcome Instrument
IRT	interactive response technology
ITT	intent-to-treat (ITT) (population)
LC-MS	liquid chromatography–mass spectrometry
LDH	lactate dehydrogenase
LS	least squares
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MELD	model for end-stage liver disease
MMRM	mixed effect model for repeated Measures
MNAR	missing not at random
PEBD	partial external biliary diversion
PedsQL	Pediatric Quality of Life Inventory
PELD	pediatric end-stage liver disease
PFIC	progressive familial intrahepatic cholestasis
PG	propylene glycol
PIC	Patient Impression of Change
PIS	Patient Impression of Severity
PK	pharmacokinetics
QD	once daily (dosing)
QTcB	QT interval corrected using Bazett's formula
QTcF	QT interval corrected using Fridericia's formula
REML	restricted maximum likelihood
SAE	serious adverse event
SAP	statistical analysis plan
sBA	serum bile acid
SOC	system organ class
SUSAR	suspected unexpected serious adverse reaction
TEAE	treatment-emergent AE
TJP2	tight junction protein 2 gene
TSB	total serum bilirubin
ULN	upper limit of normal
VAP	validation analysis plan
WHO-DD	World Health Organization Drug Dictionary

Table 1 Schedule of Assessments

Procedures[p]	Screening [a]	Dose Escalation (4-6 weeks)						Stable Dosing (20-22 weeks)										Follow-up
Visit/Subject Contact #	**Screening / V0**	**Baseline / V1**	**Subject contact**	**V2**	**Subject contact**	**V3**	**Subject contact**	**V4**	**Subject contact**	**V5**	**Subject contact**	**V6**	**Subject contact**	**V7**	**Subject contact**	**V8**	**Subject contact**	**EOT/ET V9[b]**	
Study Week	**-6**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**26**	**27**
Study Day	**Day -42**	**0**	**7**	**14**	**21**	**28**	**35**	**42**	**56**	**70**	**84**	**98**	**112**	**126**	**140**	**154**	**168**	**182**	**189**
Window (in days)	**(-5)**		**(±2)**	**(±2)**	**(±2)**	**(±2)**	**(±2)**	**(±2)**	**(±2)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±2)**
Informed Consent/Assent (10.3.1)	X																		
Eligibility Assessment	X	X																	
Demographics (7.2.4)	X																		
Medical History (7.2.3)	X																		
Physical & Vital signs[c] (7.2.5.1,7.2.5.3)	X	X		X		X		X		X		X		X		X		X	
ECG (7.2.5.4)	X	X								X								X	
Liver Ultrasound[d] (7.2.5.7)	X																	X	
Pregnancy Test[e] (7.2.5.6)	S	U				U		U		U		U		U		U		S	
Confirm PFIC Genotype Results[f] (7.1.1 and 7.2.1)	X																		
eDiary Device Provided / Returned[g]	P																	R	
ItchRO(Obs), ItchRO(Pt) & EDQ (7.2.2.1, 7.2.6.1.7)	Twice-daily completion of ItchRO & EDQ																		
Review eDiary & Assess Compliance		X		X		X		X		X		X		X		X		X	
Randomized		X																	
CIS[h] (7.2.6.1.2)		X				X		X		X		X		X		X		X	
PIS[i] (7.2.6.1.3)		X				X		X		X		X		X		X		X	
PIC (7.2.6.1.5)																		X	

Procedures[p]	Screening [a]	Dose Escalation (4-6 weeks)						Stable Dosing (20-22 weeks)											Follow-up
Visit/Subject Contact #	**Screening / V0**	**Baseline / V1**	**Subject contact**	**V2**	**Subject contact**	**V3**	**Subject contact**	**V4**	**Subject contact**	**V5**	**Subject contact**	**V6**	**Subject contact**	**V7**	**Subject contact**	**V8**	**Subject contact**	**EOT/ET V9[b]**	
Study Week	**-6**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**26**	**27**
Study Day	**Day -42**	**0**	**7**	**14**	**21**	**28**	**35**	**42**	**56**	**70**	**84**	**98**	**112**	**126**	**140**	**154**	**168**	**182**	**189**
Window (in days)	**(-5)**		**(±2)**	**(±2)**	**(±2)**	**(±2)**	**(±2)**	**(±2)**	**(±2)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±2)**
CIC (7.2.6.1.6)																		X	
PedsQL[j] (7.2.6.1.1)		X				X		X		X		X		X		X		X	
Clinician Scratch Scale (7.2.6.1.4)	X	X		X		X		X		X		X		X		X		X	
CBC with Differential (7.2.5.5, 8.3.2)	X	X		X				X		X		X		X		X		X	
Coagulation (7.2.5.5, 8.3.6)	X	X		X				X		X		X		X		X		X	
Chemistry Panel (7.2.5.5, 8.3.2, 8.3.3)	X	X		X				X		X		X		X		X		X	
Lipid Panel[k] (7.2.5.5)	X	X		X				X		X		X		X		X		X	
Urinalysis (7.2.5.5)		X						X		X				X				X	
FGF-19, Autotaxin, C4[k] (7.2.2.2, 7.2.5.5)		X						X		X				X				X	
sBA collection[k] (7.2.2.2, 7.2.5.5)	X	X		X				X		X		X		X		X		X	
Lipid Soluble Vitamins[k,l] (7.2.5.5, 8.3.5)		X						X		X		X		X		X		X	
AFP Sample (7.2.5.5)		X								X								X	
PK Sample[m] (7.2.6.2)										X								X	
Serum storage sample (7.2.5.5)		X		X				X		X		X		X		X		X	
Healthcare Utilization (7.2.6.3)				X				X		X		X		X		X		X	X
Study Medication Supplied[n] (6)		X		X		X		X		X		X		X		X			
Study Medication Administration[o] (6)		Twice-daily administration																	
Assess AEs (8)		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Prior and Concomitant Treatments (5)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

AFP=a-fetoprotein; C4=7a-hydroxy-4-cholesten-3-one; CBC=complete blood count; CIS=Caregiver Impression of Severity; CIC=Caregiver Impression of Change; ECG=electrocardiogram; EDQ=exploratory diary questionnaire; EOT=end of treatment; ET=early termination; FGF-19=fibroblast growth factor-19; ItchRO=Itch Reported Outcome; P=Provided; PedsQL=Pediatric Quality of Life Inventory; PIC=Patient Impression of Change; PIS=Patient Impression of Severity of Pruritus; PK=pharmacokinetic; R=Returned; S=Serum; U= Urine; V=visit

[a] Subjects who initially fail to meet eligibility criteria may be re-assessed during the 6-week screening period prior to being captured as a screen failure. Subjects may also be re-screened.

[b] Study sites should record dates of any future scheduled procedures related to PFIC (e.g., PEBD, ileal exclusion, liver transplant, or listed for liver transplant), if known at this visit.

[c] Blood pressure, heart rate, temperature, respiration rate. Height and weight will be measured by trained staff using standardized methodology, incl. calibrated stadiometer or headboard and calibrated balance, respectively.

[d] Screening ultrasound not required if an ultrasound or liver MRI less than 6 months old is available.

[e] Females of childbearing potential only; result must be reviewed prior to dispensing study medication.

[f] Genotyping results will be reviewed by sponsor or designee.

[g] Caregivers and age-appropriate subjects will be instructed to complete their eDiary twice daily (morning and evening).

[h] To be completed by caregivers for all subjects

[i] To be completed only by subjects aged ≥9 years at screening

[j] To be completed by subjects and caregivers using the age-appropriate PedsQL module

[k] Subjects should make every effort to fast at least 6 hours prior to collection. Water intake is permitted if necessary but not recommended.

[l] Blood samples must be drawn before administration of vitamin supplementation.

[m] PK samples will be drawn pre-dose, and approximately 2.5 hours after the morning dose.

[n] If needed, study medication may be supplied via direct-to-patient shipments in between site visits, see Section 6.4.

[o] Subjects will self-administer (or will be administered) the first dose of study medication in the clinic on Visit 1/Day 0 after breakfast under the supervision of the investigator or trained site staff.

[p] See Appendix 10 for management of study procedures during pandemic.

1 BACKGROUND INFORMATION

1.1 Indication and Current Treatment Options

Progressive familial intrahepatic cholestasis (PFIC) is a rare autosomal recessive liver disorder characterized by intrahepatic cholestasis due to canalicular bile transport defects. There are more than 4 subtypes of PFIC classified based on different mutations. PFIC1, 2, and 3 result from mutations in the ATPase phospholipid transporting 8B1 (*ATP8B1*), ATP binding cassette subfamily B member 11 (*ABCB11*), and ATP binding cassette subfamily B member 4 (*ABCB4*) genes, respectively, and all share the main clinical manifestations of cholestasis and pruritus. PFIC4 is based on mutations in the tight junction protein 2 gene (*TJP2)* and leads to failure of protein localization, disruption of tight-junction structure, and severe cholestatic liver disease (Sambrotta, 2014). In children, PFIC represents 10%–15% of causes of cholestasis and 10%–15% of indications for liver transplantations (Davit-Spraul et al., 2009). PFIC2 is the most common subtype and is diagnosed in approximately 50%–60% of PFIC patients (Al Mehaidib and Al Shahrani, 2013) while PFIC1 (also known as Byler's disease) and PFIC3 account for approximately 10%–20% (Alissa et al., 2008; Davit-Spraul et al., 2009) and 30%–40% of the PFIC population, respectively (Al Mehaidib and Al Shahrani, 2013).

PFIC is associated with early mortality, morbidity, and devastating consequences on patients' quality of life. In the absence of surgery, PFIC1 and PFIC2 are very aggressive diseases with only 10%–15% of PFIC1 and PFIC2 (depending on the variant) subjects surviving through the age of 18 (Jericho, 2015). PFIC2 is associated with a continuous progressive course of symptoms (Al Mehaidib and Al Shahrani, 2013). While extrahepatic involvement such as pancreatitis or diarrhea can be a feature of PFIC1, the initial presentation and evolution of the disease in PFIC2 tends to be more severe than in PFIC1, with persistent jaundice occurring within the first months of life and rapid progression to cirrhosis and liver failure within the first years of life (Pawlikowska et al., 2010). Interruption of the entero-hepatic circulation of bile acids through partial external biliary diversion (PEBD) surgery can lead to promising results with respect to pruritus, jaundice, and histology, both in patients with PFIC1 and PFIC2. Schukfeh, 2012 reported a dramatic 1 year outcome in patients undergoing PEBD with 13/21 (62%) of patients normalizing serum bile acids (sBAs) and liver function; however, other groups have reported overall failure rates of PEBD of up to 30% with 30%-50% of patients requiring repeat surgery (Halaweish and Chwals, 2010). Additionally, PEBD must be performed before liver fibrosis and cirrhosis are established for optimal benefit. For the vast majority of patients who do not undergo PEBD or do not respond, liver transplantation may be the only treatment option (Emerick et al., 2008; Yang et al., 2009; Arnell et al., 2010; Schukfeh et al., 2012).

As of study initiation, given the clinical outcomes associated with PFIC, including the profound negative impact on patients' and caregivers' quality of life, and the fact that there are no approved treatments, there is a clear unmet medical need for a novel treatment for this disease.

1.2 Product Background and Clinical Information

Maralixibat chloride (maralixibat; formerly SHP625, LUM001) is a potent inhibitor (IC_{50}=0.3 nM) of the apical sodium-dependent bile acid transporter/ileal bile acid transporter/solute carrier family 10 (sodium/bile acid co-transporter family) member 2 (ASBT/IBAT/SLC10A2), a transmembrane protein localized on the luminal surface of ileal enterocytes (hereafter referred to as IBAT). Maralixibat was designed to be minimally absorbed, therefore maximizing the local exposure of the molecule to the receptor and minimizing systemic exposure of the drug and limiting drug-drug interactions and systemic toxicity. These characteristics include a high molecular weight of 740 Da and the addition of a positively charged quaternary amino moiety that can interact with the negatively charged surface of the enterocyte cell membrane and prevent absorption.

The pharmacokinetics (PK), safety, and efficacy of maralixibat were assessed in a previous clinical development program that evaluated maralixibat as a potential lipid-lowering agent in patients with hypercholesterolemia. The hypercholesterolemia program was terminated by the previous sponsor for strategic and commercial reasons, and current development efforts are focused on rare pediatric cholestatic liver diseases, including Alagille syndrome (ALGS) and PFIC.

Maralixibat plasma levels (parent and potential metabolites) after chronic oral administration in adults, adolescents, and children are consistent with a drug that is minimally absorbed. Plasma concentrations, when measured, have been low to undetectable at most timepoints following oral doses at or below 20 mg in adults and in pediatric patients with ALGS receiving doses up to 280 µg/kg/day (maximum dose 20 mg/day). These results are consistent with the low absolute bioavailability (≤1%) observed in nonclinical studies with maralixibat and in a clinical study where a dose of 5 mg ^{14}C-maralixibat was administered. These data support the conclusion that maralixibat is minimally absorbed in humans. While the low bioavailability of maralixibat may complicate many classical drug development aspects such as characterization of PK (C_{max}, $t_{1/2}$, etc.), receptor occupancy, dose ranging and PK-PD correlation, etc., the potential for systemic adverse interactions with other drugs is expected to be less than for well-absorbed drugs. Drug interaction studies with both lovastatin and simvastatin and with atorvastatin indicated no clinically significant interactions with maralixibat and these statins (both parent and active metabolite).

The previous hypercholesterolemia clinical program comprised 12 studies across which a total of >1400 adult subjects were exposed to maralixibat. Maralixibat was found to be safe at investigated daily doses up to 100 mg over 28 days and up to 40 mg over 10 weeks. The most commonly reported adverse drug reactions (ADRs) were abdominal cramping/pain and diarrhea/loose stools, which are believed to be mechanism-based due to elevated bile acid concentrations in the colon. The majority of adverse events (AEs) were characterized as mild or moderate in severity, except for severe nausea, diarrhea, or abdominal pain in 6 subjects. Except for a single serious AE (SAE) of cholecystitis reported in 1 study, no other SAEs related to maralixibat have been reported in these studies and no clinically significant laboratory abnormalities were documented in maralixibat-treated subjects.

More than 180 subjects have been exposed to maralixibat, some for longer than 3 years, in the current cholestatic liver disease program, which includes 4 completed Phase 2 studies (2 studies in ALGS, 1 study in primary biliary cholangitis, and 1 study in primary sclerosing cholangitis), 1 completed Phase 1 study in overweight/obese adult male subjects, and 4 ongoing Phase 2 studies (3 studies in ALGS and 1 study in PFIC). Consistent with the previous clinical program, the most commonly reported ADRs in this program have been abdominal cramping/pain and diarrhea/loose stools, the majority of which have been mild or moderate in severity. Reported related SAEs include abdominal pain (2 cases), abdominal pain upper (3 cases), diarrhea, cholangitis, blood bilirubin increased, international normalized ratio (INR) increased, pancreatitis, alanine aminotransferase (ALT) increased, autoimmune hepatitis, hematochezia, and anemia. No deaths or important new risks have been identified to date. A program-wide review of liver safety parameters, including an independent external expert adjudication of selected cases, indicated that maralixibat may cause ALT elevations in a certain percentage (2%–5% probably, 10%–20% probably or possibly related events) of subjects with ALGS. No predictors of this treatment response have been identified so far. None of the observed events were serious and none led to liver-related morbidity or mortality. There is no evidence to date that this effect is due to a direct hepatotoxic effect from the minimal amount of maralixibat that is systemically absorbed. It is more likely that the observed abnormalities might rather be due to an indirect, compensatory effect on the liver from intestinal IBAT inhibition instead. The long-term consequences of the observed ALT elevations remain unclear. There is currently no evidence that maralixibat causes bilirubin elevations in patients with ALGS, and there is insufficient evidence of an association between bilirubin or liver enzyme elevations and maralixibat treatment in patients with PFIC.

In the ongoing Phase 2 study in subjects with PFIC (Study LUM001-501), an open-label study of the efficacy and long-term safety of maralixibat in the treatment of cholestatic liver disease in pediatric patients with PFIC, preliminary data from a first interim analysis demonstrated significant improvements in a subgroup of 6 out of 25 subjects with PFIC2 at doses of up to 280 µg/kg once daily (QD). A subsequent dose increase to 280 µg/kg twice daily (BID) produced a similar improvement in 1 additional PFIC2 subject who did not respond to the previous dose regimen. These improvements consisted of normalization or ≥70% decrease in sBA levels, concomitant with an increase in 7α-hydroxy-4-cholesten-3-one (C4) levels, a disappearance or clinically meaningful improvement in pruritus, as well as an increase in quality-of-life scores. Liver-related parameters (aspartate aminotransferase [AST]/ALT and total serum bilirubin [TSB]) were also normalized in those subjects with abnormal baseline values. Finally, the treatment responders experienced a significant catch-up growth, as indicated by a positive height and weight z-score change from baseline, as compared with those subjects who did not respond to maralixibat treatment.

As of study initiation, based on the high unmet medical need, the current absence of approved pharmacological therapy, the complications with surgical treatment options (liver transplantation and PEBD) with the need for lifelong immunosuppression or stoma care, and the overall benign safety profile and unabsorbed nature of maralixibat, the benefit-risk ratio is considered supportive of the present study. Always refer to the latest version of the maralixibat chloride Investigator's Brochure (IB) for the overall risk/benefit assessment and

the most accurate and current information regarding the drug metabolism, PK, efficacy, and safety of maralixibat.

2 STUDY OBJECTIVES AND PURPOSE

2.1 Rationale for the Study

As of study initiation, there is no approved treatment for PFIC, and available medical approaches have limited efficacy. Preliminary data from the first interim analysis from the Phase 2 Study LUM001-501 in subjects with PFIC indicate that a subgroup of subjects with PFIC2 experienced clinically significant improvements following maralixibat treatment as manifested by a large reduction or normalization of sBA, reduction in ItchRO(Obs), and normalization of TSB and ALT for those with elevated baseline values. Also, BID dosing led to a higher response rate compared with the previously used QD dosing. The current study is designed to confirm the effect of higher doses of maralixibat in PFIC2 and assess its effect on a broader PFIC population. Because all treatment responders in Study LUM001-501 were subjects with PFIC2, the primary efficacy analyses will be limited to this subpopulation.

The efficacy and safety of maralixibat in PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) and other patient subpopulations will also be assessed based on the rationale that maralixibat also has the potential to lead to sBA and pruritus reduction in the broader PFIC population and its subtypes.

2.2 Study Objectives

2.2.1 Primary Objective

- To evaluate the efficacy of maralixibat versus placebo on the **severity of pruritus** in participants with PFIC2

2.2.2 Secondary Objectives

- To evaluate the efficacy of maralixibat versus placebo on total sBA levels in participants with PFIC2
- To evaluate the efficacy of maralixibat versus placebo on the severity of pruritus in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
- To evaluate the efficacy of maralixibat versus placebo on total sBA levels in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
- To evaluate the efficacy of maralixibat versus placebo on the proportion of responders for the ItchRO(Obs) pruritus score in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)

- To evaluate the efficacy of maralixibat versus placebo on the sBA responder rate in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
- To evaluate the safety, tolerability, and pharmacokinetics of maralixibat versus placebo in all participants who receive at least 1 dose of study medication (safety population)

2.2.3 Exploratory Objectives

Additional endpoints may be explored to further characterize the efficacy and safety of maralixibat relative to placebo.

3 STUDY DESIGN

3.1 Study Design and Flow Chart

This is a 6-month international, multicenter, randomized, double-blind, placebo-controlled Phase 3 study in subjects with PFIC aged ≥12 months and <18 years at time of baseline.

The primary analysis will be conducted in the ITT population of subjects with documented biallelic mutations in ABCB11 (PFIC2), based on standard-of-care genotyping (primary cohort criteria). Subjects with a predicted complete absence of bile salt excretion pump (BSEP) function based on the type of ABCB11 mutation (PFIC2), as determined by a standard-of-care genotype, are excluded from the primary cohort and can enroll only in the supplemental cohort. Subjects with other PFIC subtypes (e.g., PFIC1/3/4, or new PFIC mutation variants) or postsurgical subjects (e.g., after internal or external biliary diversion surgery or subjects with reversal of biliary diversion surgery) will be enrolled in a separate supplemental cohort and evaluated as part of secondary and exploratory analyses (see Section 9.8).

Subjects in the primary cohort will be randomized in a 1:1 ratio to receive maralixibat or placebo. Subjects in the supplemental cohort will be randomized in a 1:1 ratio to receive maralixibat or placebo within the following subcohorts: a) PFIC1; b) PFIC3; c) all other supplemental cohort subjects. For further information about the treatment allocation, refer to Section 6.2.2. All subjects will receive study treatment in addition to standard-of-care therapy, including antipruritic concomitant treatment (for prohibited treatments, see Section 5.2.2).

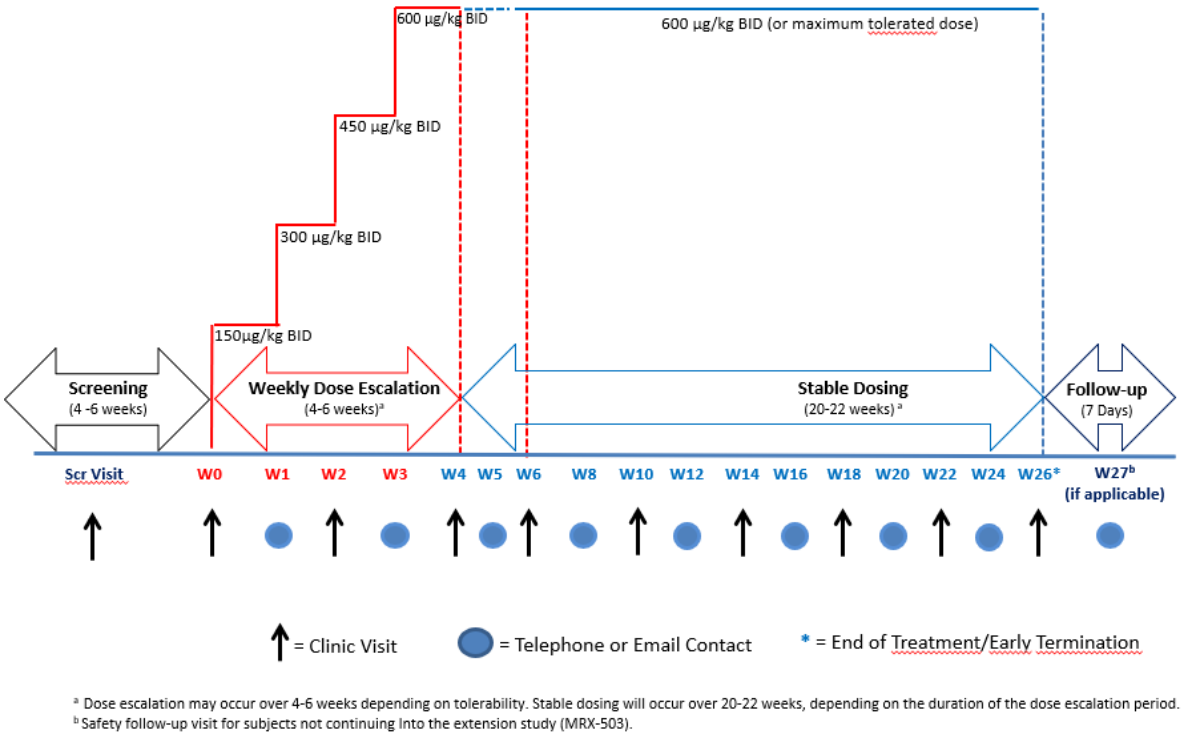
In the primary cohort, approximately 30 subjects will be enrolled in order to achieve 26 subjects (13 per treatment group) completing the study. In the supplemental cohort, overall enrollment will be capped at a maximum of 60 subjects.

The study will be divided into 4 parts (see Figure 1):

1. Screening period (up to 6 weeks)
2. Dose Escalation period (4–6 weeks)

3. Stable Dosing period (20–22 weeks)
4. Safety Follow-up period (7 days; for subjects discontinuing early and for subjects not enrolling into the extension Study MRX-503; see Section 7.1.4)

Figure 1 Study Design Flow Chart



BID=twice daily; Scr=screening; V=visit; W=week

For information about the schedule of assessments throughout the study, refer to Table 1 and Section 7. For information on dosing, including dose escalation requirements, refer to Section 6.2.3.

3.1.1 Rationale for Treatment and Dose Escalation

The lower age limit for enrollment in this study is 1 year, based on the fact that the gastrointestinal (GI) tract is fully mature at 1 year in humans and higher plasma exposures were observed in rats with immature GI tracts in juvenile toxicokinetic studies. Given the extremely low absorption of maralixibat in adults, minimal exposure to plasma is anticipated in eligible pediatric patients, so the pediatric doses have been selected based upon body surface area scaling of doses shown to be safe and tolerable in adults in a Phase 1 healthy volunteer study. In Study SHP625-101, a blinded, placebo-controlled, randomized, multiple-dose study, bile acids levels in feces increased with escalating doses up to 100 mg QD and 50 mg BID of maralixibat, without any meaningful changes in the safety or tolerability profile. Thus, higher doses of maralixibat might be more efficacious, whereas a

BID regimen has the potential to allow for more complete coverage of the distal ileum throughout the day. Of note, these doses correspond to approximately 1400 µg/kg QD and 700µg/kg BID in a 70-kg adult, respectively.

Ideally, dosing in pediatric subjects should be scaled from that in adults based on intestinal length—that is, milligram of drug per centimeter of intestine. Differing relationships between intestinal mucosal surface area, age, and body weight have been reported, including data indicating that the average length of the small intestine increases with age from birth through 20 years; this relationship followed a curve that is similar to the height and weight growth curves (Weaver et al. 1991). However, a plateau had not been reached at the maximum age examined (20 years), precluding predictions of intestinal length for older adults and thus scaling to infants and children based on estimated intestinal length. An analysis of intestinal length as a function of age, weight, and height in adult cadavers was conducted by Hounnou et al. (2002). Their analysis demonstrated that age had a negative correlation and body weight a positive correlation with intestinal length. Taken as a whole, the existing analyses are inconclusive with respect to the dependent variables that predict intestinal length. Consequently, the most reasonable approach is to calculate doses in a pediatric subject from those in adults using a direct mg/kg scaling. For reference, in an average adult subject weighing 70 kg, the proposed maximum target dose of 600 µg/kg BID corresponds to a total daily dose of 42 mg BID. Doses up to 400 µg/kg BID are generally well tolerated in the currently ongoing maralixibat Phase 2 studies in children with PFIC and ALGS.

Sample daily exposure (mg-day) across proposed dose levels for subjects ranging in weight from 5–70 kg is provided in Table 2.

Table 2 Sample Daily Exposure (mg/day) in Pediatric Subjects

Weight (kg)	Daily Exposure of Maralixibat (mg)			
	Dose Level 1 (150 µg/kg - BID)	**Dose Level 2 (300 µg/kg - BID)**	**Dose Level 3 (450 µg/kg - BID)**	**Dose Level 4 (600 µg/kg - BID)**
5	1.5	3.0	4.5	6.0
10	3.0	6.0	9.0	12.0
20	6.0	12.0	18.0	24.0
30	9.0	18.0	27.0	36.0
40	12.0	24.0	36.0	48.0
50	15.0	30.0	45.0	60.0
60	18.0	36.0	54.0	72.0
70	21.0	42.0	63.0	84.0

3.1.2 Rationale for Primary Endpoint

The primary endpoint is the mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 to 26. The selection of ItchRO(Obs) as the primary endpoint of therapeutic response to maralixibat treatment was based on the clinically significant

improvements observed in 7 subjects in the ongoing Phase 2 PFIC Study LUM001-501. Initial efficacy data collected from the first interim data analysis demonstrated signals of broad, significant, and sustained efficacy in a subgroup of 7 subjects, all with PFIC2. Among other positive clinical outcomes, these subjects experienced large reductions in pruritus, as measured by improvements in ItchRO(Obs) score of ≥1.0. These clinically significant reductions were sustained through 48 weeks of treatment for most subjects, showed little variability throughout the study, and closely mirrored the results frequently observed after PEBD (Yang et al. 2009; Stapelbroek et al. 2010). Full treatment effect is usually observed by 12 weeks of treatment. A seventh subject was identified as exhibiting a partial response that was characterized by significant improvements in sBA (reduction of 88% from baseline) and in ItchRO(Obs) scores. Both measures of efficacy remained largely sustained for the duration of the study for all these subjects, except during intercurrent illnesses (e.g., GI or upper respiratory infections). This pattern of consistent, monotonic, and concomitant improvements across multiple parameters in these subjects is consistent with what has previously been described following biliary diversion surgery and inconsistent with the usual natural history of PFIC, which is characterized by progressive deterioration in liver function without periods of normalization of pruritus, bile acids, or liver enzymes.

3.1.3 Rationale for Study Population

Limiting the primary analysis to subjects with PFIC2 is based on the following:
1. All responders in Study LUM001-501 are subjects with PFIC2; 2. PFIC2 patients most commonly lack extrahepatic abnormalities; and 3. PFIC2 represents the most prevalent subgroup of PFIC patients and, therefore, constitutes the biggest unmet need.

3.2 Duration and Study Completion Definition

The subject's maximum duration of participation is expected to be up to 33 weeks.

The Study Completion Date is defined as the date the final subject, across all sites, completes his or her final protocol-defined assessment. Note that this includes the follow-up visit or contact, whichever is later (refer to Section 7.1.4 for the defined follow-up period for this protocol) but not the long-term clinical follow-up of persistent safety findings.

At the end of the 26-week double-blind treatment period, subjects will have the opportunity to enroll into the long-term, extension study (MRX-503). Subjects who do not enroll into the treatment extension study or who discontinue at any time during this study will have a final safety follow-up telephone call 7 days after the last dose of study medication (see Section 7.1.4). Subjects who discontinue from the study for safety reasons are followed as long as clinically indicated or until no further improvement with an adverse outcome is expected.

3.3 Sites and Regions

The study will be conducted in approximately 40 sites in North America, Europe, Asia, and South America. Other regions may be added, if necessary, for enrollment purposes.

4 STUDY POPULATION

4.1 Inclusion Criteria

A subject will not be considered eligible for the study without meeting all the criteria below.

1. Informed consent and assent (as applicable) per Institutional Review Board/Ethics Committee (IRB/EC)
2. Male or female subjects with a body weight ≥5.0 kg who are ≥12 months and <18 years of age at time of baseline
3. Cholestasis as manifested by total sBA ≥3× ULN (applies to primary cohort only)
4. An average AM ItchRO(Obs) score ≥1.5 during 4 consecutive weeks of the screening period, leading to the baseline visit (Visit 1)
5. Completion of at least 21 valid* morning ItchRO(Obs) entries during 4 consecutive weeks of the screening period, leading to the baseline visit (Visit 1) (*valid=completed and not answered as "I don't know"; maximum allowed invalid reports=7, no more than 2 invalid reports during the last 7 days before randomization)
6. Diagnosis of PFIC based on:

   Chronic cholestasis as manifested by persistent (>6 months) pruritus in addition to biochemical abnormalities and/or pathological evidence of progressive liver disease

   and

   Primary Cohort:

   Subjects with genetic testing results consistent with biallelic disease-causing variation in ABCB11 (PFIC2), based on standard-of-care genotyping

   Supplemental Cohort:

   i. Subjects with genetic testing results consistent with biallelic disease-causing variation in ATP8B1 (PFIC1), ABCB4 (PFIC3), or TJP2 (PFIC4), based on standard-of-care genotyping

   ii. Subjects with PFIC phenotype without a known mutation or with another known mutation not described above or with intermittent cholestasis as manifested by fluctuating sBA levels

   iii. Subjects with PFIC after internal or external biliary diversion surgery or for whom internal or external biliary diversion surgery was reversed
7. Male and females of non-childbearing potential. Males and non-pregnant, non-lactating females of childbearing potential who are sexually active must agree to use acceptable methods of contraception during the study and 30 days following the last dose of the study medication. Females of childbearing potential must have a negative pregnancy test result.

8. Access to email or telephone for scheduled remote visits
9. Ability to read and understand the questionnaires (both caregivers and subjects above the age of assent)
10. Access to consistent caregiver(s) during the study
11. Subject and caregiver willingness to comply with all study visits and requirements

4.2 Exclusion Criteria

A subject will be excluded from the study if any of the following exclusion criteria are met:

1. Predicted complete absence of bile salt excretion pump (BSEP) function based on the type of ABCB11 mutation (PFIC2), as determined by a standard-of-care genotyping (applies to primary cohort only). Subjects can enter the study in the Supplemental Cohort (under inclusion criteria 6.ii or 6.iii).
2. Recurrent intrahepatic cholestasis, indicated by a history of sBA levels <3x ULN or intermittent pruritus (applies to primary cohort only)
3. Current or recent history (<1 year) of atopic dermatitis or other non-cholestatic diseases associated with pruritus
4. History of surgical disruption of the enterohepatic circulation (applies to primary cohort only)
5. Chronic diarrhea requiring intravenous fluid or nutritional intervention for the diarrhea and/or its sequelae at screening or during the 6 months prior to screening
6. Previous or need for imminent liver transplant
7. Decompensated cirrhosis (international normalized ratio [INR] >1.5, and/or albumin <30 g/L, history or presence of clinically significant ascites, and/or variceal hemorrhage, and/or encephalopathy)
8. ALT or TSB >15× ULN at screening
9. Presence of other liver disease
10. Presence of any other disease or condition known to interfere with the absorption, distribution, metabolism, or excretion of drugs, including bile salt metabolism in the intestine (e.g., inflammatory bowel disease), per investigator discretion
11. Possible malignant liver mass on imaging, including screening ultrasound
12. Known diagnosis of human immunodeficiency virus (HIV) infection
13. Any prior cancer diagnosis (except for in situ carcinoma) within 5 years before the screening visit (Visit 0)
14. Any known history of alcohol or substance abuse
15. Administration of bile acids or lipid binding resins or phenylbutyrates during the screening period
16. Criterion has been deleted as of Amendment 3.

17. Administration of any investigational drug, biologic, or medical device during the screening period
18. Previous use of an ileal bile acid transporter inhibitor (IBATi)
19. History of nonadherence to medical regimens, unreliability, medical condition, mental instability or cognitive impairment that, in the opinion of the investigator or sponsor medical monitor, could compromise the validity of informed consent, compromise the safety of the subject, or lead to nonadherence with the study protocol or inability to conduct the study procedures.
20. Known hypersensitivity to maralixibat or any of its excipients

4.3 Reproductive Potential

Local, country-specific guidelines should always be followed where there is regional variation for which methods of contraception are considered acceptable.

4.3.1 Female Contraception

- Females are considered to be of non-childbearing potential if they are premenarchal and either Tanner Stage 1 or younger than age 9 years.
- Females of childbearing potential must have a negative β-hCG pregnancy test result as outlined in Table 1. Results of pregnancy test must be reviewed prior to dispensing study medication.
- Females of childbearing potential must agree to abstain from sexual activity that could result in pregnancy or, if they are sexually active or become sexually active during the study, agree to use acceptable methods of contraception, as defined below, throughout the study period and for 30 days following the last dose of study medication. If hormonal contraceptives are used, they should be administered according to the package insert.

Acceptable methods of contraception are:

- Intrauterine devices plus condoms
- Double-barrier methods (e.g., condoms and diaphragms with spermicidal gel or foam)
- Hormonal contraceptives (oral, depot, patch, injectable, or vaginal ring), stabilized for at least 30 days prior to the screening visit (Visit 0), plus condoms. Note: If subject becomes sexually active during the study, they should use one of the other acceptable methods noted above in addition to the hormonal contraceptive until it has been stabilized for 30 days.

4.3.2 Male Contraception

Contraception is required for all sexually active male subjects and their partners. All male subjects must agree not to donate sperm, and to use 1 of the following approved methods of contraception until 30 days following the last dose of study medication:

- Male condom with spermicide
- Intrauterine device with spermicide (use by female sexual partner)
- Female condom with spermicide (use by female sexual partner)
- Contraceptive sponge with spermicide (use by female sexual partner)
- Intravaginal system (e.g., vaginal ring with spermicide, a diaphragm with spermicide, or a cervical cap with spermicide) (use by female sexual partner)
- Oral, implantable, transdermal, or injectable hormonal contraceptive (use by female sexual partner).

4.4 Withdrawal/Discontinuation from Study

A subject may withdraw from the study at any time for any reason without prejudice to future medical care by the physician or at the institution. The investigator or sponsor may discontinue the subject from the study at any time (refer to Section 4.4.1). The investigator is encouraged to discuss discontinuation of a subject with the medical monitor when possible. The reason for termination and the date of stopping study medication must be recorded in source documents. The evaluations listed for the Week 26 visit (Visit 9) are to be performed as completely as possible for any subject who discontinues early (see Section 7.1.3.3).

Subjects who discontinue from the study for safety reasons are followed as long as clinically indicated or until no further improvement with an adverse outcome is expected.

Subjects who discontinue from the study will not be replaced.

4.4.1 Reasons for Discontinuation from Study

Medically important events that in the opinion of the investigator or medical monitor would compromise the subject's ability to safely continue the study would warrant the subject's discontinuation from the study. The reason for discontinuation must be recorded in the subject's source document. If a subject is discontinued for more than 1 reason, each reason should be documented in the source and the most clinically relevant reason should be indicated.

Reasons for discontinuation include but are not limited to:

- Death
- Noncompliance with study procedures
- Adverse event
- Withdrawal by subject/parent/guardian
- Physician decision
- Loss to follow-up
- Pregnancy
- Imminent need for liver transplant
- Initiation or increase in dose of anti-pruritic or prohibited medication (see Section 5.2.2)
- Progressive disease (defined as progression of underlying PFIC which, in the opinion of the investigator, is due to the disease itself and can be explained by the natural history of PFIC—i.e., worsening of pruritus, cholestasis, or liver disease) and not due to medications or other factors.
- Other

Additional reasons may be referenced on the Case Report Form (CRF).

If subjects discontinue from the study due to disease progression and require rescue therapy with either PEBD or listing for liver transplant, both the reason for discontinuation and outcome (i.e., PEBD, cholecystostomy tube, ileal exclusion, liver transplant, or listed for liver transplant) should be recorded in the subject's source document. If known, the date of the future scheduled procedure should also be recorded. The information regarding PEBD, liver transplant, or listing for transplant will be collected in the CRF at the time the subject completes/early terminates from the study (Visit 9; EOT/ET) until the completion of the follow-up period (Week 27).

4.4.2 Subjects "Lost to Follow-up" Prior to Last Scheduled Visit

A minimum of 3 documented attempts must be made to contact any subject lost to follow-up at any time point prior to the last scheduled contact (office visit or telephone contact). At least 1 of these documented attempts must include a written communication sent to the subject's last known address via courier or mail (with an acknowledgement of receipt request) asking that the subject return to the site for final safety evaluations and return any unused study medication.

5 PRIOR AND CONCOMITANT TREATMENT

5.1 Prior Treatment

Specific prior concomitant treatments and therapies administered for underlying liver disease, treatment of pruritus, or treatment of cholestasis will be collected. Prior treatment information must be recorded in the subject's source document for data collection.

5.2 Concomitant Treatment

A subject should not be instructed to discontinue use of any medication or treatment to participate in this study until informed consent (or assent as applicable) has been obtained. Concomitant treatment refers to all treatment, including concomitant therapies, as well as herbal treatments, vitamins, behavioral treatment, non-pharmacological treatment, such as psychotherapy, taken between the dates of the first dose of study medication and the end of the subject's participation in the study (Week 26 or Week 27), inclusive. Concomitant treatment information must be recorded in the subject's source document. Investigators must ensure that subjects receive lipid soluble vitamin supplements, as needed, for the duration of the study.

5.2.1 Permitted Treatment

The following medications to treat cholestasis/pruritus are allowed during the study if the subject has been on a stable dosing regimen (i.e., same dose and frequency in the 30 days prior to the screening visit [Visit 0]) and will continue this dosing regimen throughout study participation. The investigator must contact the medical monitor to discuss any changes to and introduction of concomitant treatments to treat cholestasis/pruritus not listed here that could impact the outcome of the study. Dose adaptations to body weight changes are allowed.

- Ursodeoxycholic or cholic acid
- Rifampicin
- Diphenhydramine
- Hydroxyzine
- Naltrexone
- Naloxone
- Phenobarbital
- Antihistamines

5.2.2 Prohibited Treatment

Table 3 details the prohibited treatments and the time restriction period prior to Baseline (Visit 1/Day 0). None of the medications listed in Table 3 are allowed during the conduct of the study.

Table 3 Prohibited Treatments and Time Restriction Prior to Baseline

Prohibited Treatment	Time Restriction prior to Baseline (Visit 1/Day 0)
Bile acid or lipid binding resins	Any time during the screening period
Phenylbutyrates	Any time during the screening period
Any investigational drug, biologic, or medical device	Any time during the screening period
Another IBATi/ASBTi	Any time before the study

IBATi=ileal bile acid transporter inhibitor; ASBTi=apical sodium-dependent bile acid transporter inhibitor

6 STUDY MEDICATION

6.1 Identity of Investigational Product

The IP is maralixibat chloride (maralixibat; formerly SHP625, LUM001), which will be provided as an oral solution form (e.g., 5, 10, 15, and 20 mg/mL) along with either 0.5, 1.0, or 3.0-mL sized dosing dispensers. The reference/comparator product is a placebo which will also be provided as an oral solution along with dosing dispensers of the same size (either 0.5 mL, 1.0 mL, or 3.0 mL). Maralixibat and placebo (the study medication) are presented in 30-mL volumes packaged in 30-mL size amber colored PET bottles and require refrigerated storage conditions (2°C–8°C).

One of the excipients in the maralixibat oral solution is propylene glycol (PG); in order to limit subject exposure to PG, a specific strength of oral solution will be prescribed to a given subject based on body weight and target dose. The dosing plan will limit PG exposure to ≤26 mg/kg/day and, at the same time, will provide reasonable (not too high or too low) dosing volumes to ensure accurate dosing.

Additional information is provided in the maralixibat chloride IB and the pharmacy manual.

6.1.1 Blinding the Treatment Assignment

The placebo solution will contain all components of the IP except the active drug substance. All packaged study medication components, including the dosing dispensers, will be identical in order to maintain the blind.

6.2 Administration of Study Medication

6.2.1 Interactive Response Technology for Study Medication Management

An interactive response technology (IRT) will be used for screening and enrolling subjects, randomization, study medication supply dispensation and management, inventory management and supply ordering, study medication expiration tracking and management, and emergency unblinding. Please refer to the study manual for additional details regarding the IRT.

The investigator or designee will access the IRT at the screening visit (Visit 0) to record subject-specific information (i.e., unique subject number, age [years and months], etc.).

At the baseline visit (Visit 1), the investigator or designee will again access the IRT to either document a screen failure or, if the subject has met all entry criteria, to randomize the subject. For randomized subjects, the IRT will provide one or more bottle identification number(s) to dispense for treatment.

A user manual with specific functions and instructions for the IRT will be provided to the site, and site personnel will receive training.

6.2.2 Allocation of Treatment to Subjects

Individual subject treatment is automatically assigned by the IRT. Subjects will be randomized after confirmation of study eligibility in a 1:1 ratio via a computer-generated randomization schedule to receive maralixibat or placebo, stratified by cohort (refer to Section 3.1). The subject's randomization number represents a unique number corresponding to the treatment allocated to the subject.

Siblings will be randomized in a blinded manner to the same treatment group. Additional details are included within the SAP.

6.2.3 Dosing

Subjects will be administered (or will self-administer) varying volumes of ready-to-use oral solution study medication at each dosing visit, starting with the baseline visit (Visit 1). The dosing volume is determined based on the individual body weight, the dose level according to the dose escalation plan (150, 300, 450, or 600 µg/kg), and the strength of solution being administered (5, 10, 15, or 20 mg/mL).

At each dosing visit, the investigator or designee will access the IRT, and enter the subject's weight and information on dose escalation status. This information will be used by the IRT to identify the appropriate strength of solution and dosing volume, and to assign a corresponding bottle number(s) to be dispensed to the subject. Additional details will be provided in the pharmacy manual.

Subjects will take the first dose of study medication at the baseline visit (Visit 1) under the supervision of the investigator or trained staff.

Study medication administration will take place during the Dose Escalation and Stable Dosing periods of the study based on a BID regimen. The morning dose should be administered approximately 30 minutes before breakfast and the evening dose approximately 30 minutes before the main evening meal. The doses will be administered prior to meals rather than q12h in order to better cover the luminal bile acid release associated with meals. Study medication should be administered approximately at the same time each day throughout the study. The interval between administration of study medication and any concomitant treatment should not change during the study. The morning doses administered at the site during the site visits should be taken from the new bottles allocated during that respective visit and not from the bottles returned by the subject.

For subjects assigned to maralixibat, the Dose Escalation period will consist of the following weekly steps:

- Dose level 1: 150 µg/kg maralixibat BID for 1 week
- Dose level 2: 300 µg/kg maralixibat BID for 1 week
- Dose level 3: 450 µg/kg maralixibat BID for 1 week
- Dose level 4: 600 µg/kg maralixibat BID for the remaining duration of the study

Dose escalation should occur in the absence of major safety (e.g., liver parameters) or tolerability (e.g., GI-related TEAEs) concerns related or possibly related to study medication. Subjects with such safety concerns can be down-titrated to a lower, previously tolerated dose level for 1 week before continuing dose escalation. The minimum dose to continue in the study will be 150 µg/kg maralixibat BID; subjects who cannot tolerate this dose will be discontinued from the study. Investigators should also review study medication compliance during the Dose Escalation period to ensure that subjects have adequate exposure to study medication to assess safety and tolerability. Any compliance concerns should be discussed with the medical monitor.

Investigators have up to Week 6 to determine the maximum tolerated dose; if re-challenges or further dose escalations fail, the subject will remain on the maximum tolerated dose level for the remainder of the Stable Dosing period to complete a minimum of 20 weeks of stable dosing.

During the Stable Dosing period, subjects will remain on the maximum tolerated dose level for the remainder of the study. Dose reductions are allowed for safety or tolerability reasons down to a minimum level of 150 µg/kg BID. Subjects who cannot tolerate this dose will be discontinued from the study.

6.2.4 Unblinding the Treatment Assignment

The treatment assignment must not be unblinded during the study except for expedited safety reporting and in emergency situations where the identification of the study medication is required for further treatment of the subject. The investigator should make an effort to

contact the medical monitor before emergency unblinding occurs or as soon as possible after the investigator has been unblinded without revealing the treatment assignment to the medical monitor. In any event, the medical monitor and sponsor must be informed about the code break as soon as possible.

In the event that the treatment assignment is unblinded, the date and the signature of the person who was unblinded, and the reason for unblinding are recorded in the source documents. Upon breaking the blind in emergency situations, the subject will be withdrawn from the study, but should be followed up for safety purposes. Unblinding will be processed through the IRT vendor, and instructions will be available in the user manual.

Except for samples collected at the screening and baseline visits, sBA results will remain blinded for all sites and the blinded study team until after database lock.

6.3 Labeling, Packaging, Storage, and Handling

6.3.1 Labeling

All study medication is labeled according to country-specific requirements. Additional labels may, on a case-by-case basis, be applied to the study medication in order to satisfy local or institutional requirements, but must not:

- Contradict the clinical study label.
- Obscure the clinical study label.
- Identify the study subject by name.

Additional labels may not be added without the sponsor's prior full agreement.

6.3.2 Packaging

Study medication is packaged in the following labeled containers:

- Ready-to-use oral solution of maralixibat 5 mg/mL
- Ready-to-use oral solution of maralixibat 10 mg/mL
- Ready-to-use oral solution of maralixibat 15 mg/mL
- Ready-to-use oral solution of maralixibat 20 mg/mL
- Maralixibat matching placebo

The sponsor or designee will provide 0.5, 1.0, and 3.0 mL sized oral dosing dispensers, bottle adapters, and Child Resistant Caps to clinical sites. Changes to sponsor-supplied packaging may not occur without full agreement in advance by the sponsor.

6.3.3 Storage

The investigator has overall responsibility for ensuring that study medication is stored in a secure, limited-access location at refrigerated storage conditions (2°C–8°C). Temperature monitoring is required at the storage location to ensure that the study medication is maintained within an established temperature range. The sponsor must be notified immediately upon discovery of any excursion from the established range, or if there are any changes to the storage area of the study medication that could affect the integrity of the study medication.

6.4 Drug Accountability

Investigators will be provided with sufficient amounts of the study medication to carry out this protocol for the agreed number of subjects. The investigator or designee will acknowledge receipt of the study medication, documenting shipment content and condition. Accurate records of all study medication dispensed, used, returned, and/or destroyed must be maintained.

The investigator has overall responsibility for administering/dispensing study medication. Where permissible, tasks may be delegated to a qualified designee (e.g., a pharmacist) who is adequately trained in the protocol and who works under the direct supervision of the investigator. Study medication will be dispensed at baseline and each visit in quantities sufficient for dosing until the next scheduled visit. If necessary, direct from site to patient shipment of medication is allowed between visits.

The investigator or his/her designee will dispense the study medication only to subjects included in this study following the procedures set out in the study protocol. Each subject will be given only the study medication assigned. All dispensed medication will be documented in IRT.

The investigator is responsible for ensuring the retrieval of all study supplies from subjects. Subjects must be instructed to bring their unused and empty/used study medication to every visit. Drug accountability must be assessed at the bottle level. The pharmacist/delegated person will record details on the drug accountability form or similar.

Other than study medication dispensed to subjects, no stock or returned inventory may be removed from the site where originally shipped without prior knowledge and written consent by the sponsor. If such transfer is authorized by the sponsor, all applicable local, state, and national laws must be adhered to for the transfer.

The sponsor or its representatives must be permitted access to review the supplies storage, distribution procedures and records provided that the blind of the study is not compromised.

At the end of the study, or as instructed by the sponsor, all unused stock, subject-returned study medication, and empty/used study medication packaging are to be returned or destroyed. Study medication must be counted and verified by clinical site personnel and the sponsor (or designated CRO) prior to return or destruction. Shipment return or destruction

forms must be completed prior to shipment from the site or destruction. Shipment of all returned study medication must comply with local, state, and national laws.

Based on entries in the site drug accountability forms, it must be possible to reconcile study medication delivered with those used and returned. All study medication must be accounted for, and all discrepancies investigated and documented to the sponsor's satisfaction.

6.5 Subject Compliance

Subject compliance with study procedures and treatment compliance will be assessed at each treatment visit by the site staff.

During the Dose Escalation period, subject compliance will be reviewed as part of the dose escalation assessment. Any concerns with compliance should be discussed with the medical monitor.

During the Stable Dosing period, subjects will be allowed a maximum of 28 days (both morning and afternoon dose) of cumulative dose interruption. During the last 12 weeks of the Stable Dosing period (Week 15–26), subjects will be allowed a maximum of 17 days (both morning and afternoon dose) of cumulative dose interruption, but not more than 7 consecutive days. Any dose interruption longer than the maximum allowed will be documented as a protocol deviation. Subjects may remain on study during dose interruptions and should continue to complete all regularly scheduled subject study visits and assessments as outlined in Table 1. After any dose interruption, subjects will reinitiate dosing at their maximum tolerated dose.

Any subject who has over 56 days (both morning and afternoon dose) of cumulative dose interruption during the Stable Dosing period should be discontinued from the study. Procedures outlined in Section 7.1.3.3 should be followed for any subject who is discontinued early from the study.

7 STUDY PROCEDURES

See Appendix 10 for management of study procedures during pandemic.

7.1 Study Schedule

The study procedures and assessments to be performed throughout the study can be found in the schedule of assessments in Table 1. Further instructions about study assessments are provided in this section and in Section 7.2.

Prior to performing any study-related procedures (including those related to screening), the investigator or his/her designee must obtain written informed consent (or assent as applicable) from the subject (as per local requirements). For information about informed consent process see Section 10.3.1.

Subject numbers are assigned to all subjects as they consent to take part in the study. Within each site, the subject number is assigned to subjects according to the sequence of presentation for study participation.

Where possible and permitted, study visits can be organized through qualified health care professionals to perform study assessments at the subject's home (or other agreed upon location), once approved by the Principal Investigator and medical monitor.

7.1.1 Screening Period (Day -42 to Day 0)

The screening period starts when informed consent (or assent as applicable) is signed. The duration of the screening period is up to 6 weeks during which all procedures listed for the screening visit (Visit 0) in Table 1 must be completed. The screening period will last at a minimum 4 weeks to confirm the morning ItchRO(Obs) score ≥1.5 and will allow for the evaluation of each subject's eligibility for inclusion into the study.

Subjects who meet eligibility criteria after completion of all screening visit (Visit 0) assessments will enter the 26-week double-blind treatment period. This period should not commence until all screening assessments required to confirm eligibility for randomization have been completed and the PFIC genotype results have been confirmed by sponsor or designee.

If the subject does not meet eligibility criteria following completion of screening assessments, the investigator or designee will document the subject as a screen failure in the IRT. A screen failure is a subject who has given informed consent and failed to meet the inclusion and/or met at least 1 of the exclusion criteria and has not been randomized or administered study medication.

In exceptional cases, subjects who initially fail to meet all eligibility criteria may be re-screened at a later timepoint, after consultation with the medical monitor.

Screen failure subject numbers cannot be reassigned to randomized subjects.

A 5-day visit window will be permitted in the case the 42-day screening period needs to be extended. This window will allow for blood sample retesting at the investigator's discretion.

7.1.2 Dose Escalation Period: Visit 1 (Week 0) to Visits 3–4 (Week 4–6)

Subjects who continue to meet all eligibility criteria will be randomized at the baseline visit (Visit 1) 1:1 to receive either maralixibat or placebo as described in Section 6.2.3. The investigator or assigned site staff will access the IRT to randomize the subject and dispense the study medication (Section 6.2.1).

The double-blind Dose Escalation treatment period will comprise of 4–6 weeks (maximum 6 weeks). All assessments and procedures listed for Visits 1-2 and interim contacts (4-week Dose Escalation) or Visits 1-3 and interim contacts (6-week Dose Escalation), in Table 1 should be completed during this period. Subject contacts may be by telephone or email.

During this period, a ±2-day visit window will be permitted, unless otherwise specified. Visit dates and acceptable windows are calculated from the baseline visit (Visit 1/Day 0).

7.1.3 Stable Dosing Period

7.1.3.1 Visits 3–8 (Weeks 4, 6, 10, 14, 18, and 22)

Subjects will return to the site for Visits 3–8, or Visits 4–8 for subjects requiring 6-week dose escalations. Assessments and procedures will be performed as outlined in Table 1.

For Visits 3–4, a ±2-day visit window will be permitted. For Visits 5–8, a ±3-day visit window will be permitted. Visit dates and acceptable windows are calculated from the baseline visit (Visit 1/Day 0).

7.1.3.2 Subject Contact (Weeks 5, 8, 12, 16, 20, and 24)

Subjects/caregivers will receive an email or a telephone call on Weeks 5, 8, 12, 16, 20, and 24 during which the use of concomitant treatments and AEs will be reviewed as outlined in Table 1. For subjects requiring 6 weeks of dose escalation, the first subject contact during the Stable Dosing period will occur at Week 8.

For subject contacts through Week 8, a ±2-day visit window will be permitted. For subject contacts from Week 12 through 24, a ±3-day visit window will be permitted. Visit dates and acceptable windows are calculated from the baseline visit (Visit 1/Day 0).

7.1.3.3 Visit 9 [End of Treatment (EOT)/Early Termination (ET)] (Week 26)

Subjects will return to the site for Visit 9 at the end of treatment (EOT) or upon early termination (ET) and the procedures and assessments outlined in Table 1 will be performed.

If study medication is permanently discontinued early, regardless of the reason, the subject will be discontinued from the study (refer to Section 4.4), and the evaluations listed for the Week 26 (Visit 9/ET) are to be performed. The subject should undergo the ET assessments at the visit during which study medication was discontinued, or the subject should be scheduled for an additional ET visit as soon as possible, if study medication was discontinued between visits.

Subjects who discontinue prematurely at any time during the study will also have a final safety follow-up telephone call 7±2 days after the last dose of study medication. See Section 7.1.4. At the EOT/ET Visit, subjects who discontinue prematurely and subjects who will not participate in the MRX-503 extension study will be asked if they are willing to provide informed consent and assent (as appropriate) for the investigator and/or designee to contact them, their caregivers and/or their family doctor/primary care physician to obtain information on how their disease has progressed and whether or not they have had PEBD surgery or have been listed for, or had a liver transplant or any other procedure relevant to the management of their cholestatic liver disease (i.e., PEBD, cholecystostomy tube, ileal exclusion, liver transplant, or listed for liver transplant, other). Any increase in concomitant

medications or introduction of new medication for the treatment of pruritus and/or cholestatic liver disease will also be collected. The follow-up will be done at least annually (from the date of the EOT/ET Visit), documented and entered in the CRF accordingly, until disease progression or an event occurs (whichever comes first).

The subject should be asked prior to Visit 9/EOT if he or she would like to participate in the MRX-503 extension study, because the Baseline visit of MRX-503 should occur on the same day as Visit 9/EOT in this study.

7.1.4 Safety Follow-up (Day 7 Post-treatment)

The safety follow-up period for this protocol is 7 days from the last dose of study medication. Subjects/caregivers who do not enroll into the extension study (MRX-503) will have a final safety follow-up telephone call at Week 27.

Subjects who discontinue prematurely at any time during the study will have a final safety follow-up telephone call 7 days after the last dose of study medication. Subjects who discontinue from the study for safety reasons are also followed-up as long as clinically indicated or until no further improvement with an adverse outcome is expected. All AEs and SAEs that are not resolved at the time of this contact will be followed to closure by the site (see Section 8.1).

7.1.5 Additional Care of Subjects After the Study

No aftercare is planned for this study for those subjects who do not enroll in the extension study (MRX-503).

7.2 Study Evaluations and Procedures

7.2.1 Genetic Testing Results

ATP8B1 (PFIC1), *ABCB11* (PFIC2), *ABCB4* (PFIC3), *TJP2* (PFIC4), NR1H4 (PFIC5), and MYO5B (PFIC6) mutations are predictive of PFIC. Standard of care genotyping results will be reviewed and documented by the sponsor or designee for confirmation of PFIC subtype and for determination of cohort assignment.

7.2.2 Efficacy

7.2.2.1 Itch Reported Outcome (ItchRO™)

Pruritus will be assessed using the Itch caregiver/patient reported outcome measure (ItchRO™) administered as a twice daily electronic diary as described in Appendix 2 and Appendix 3. Caregivers for all subjects will complete the Observer instrument: ItchRO(Obs). The ItchRO(Obs) should be completed by the same caregiver for consistency, whenever possible.

The ItchRO(Obs) eligibility must be determined after a minimum of 4 weeks of ItchRO(Obs) twice-daily diary entry. If the subject does not meet eligibility after 4 weeks, ItchRO(Obs) eligibility should be re-assessed at week 5, and also at week 6 if eligibility is not met at week 5. A maximum of 3 attempts can be made to determine ItchRO(Obs) eligibility during the screening period.

Only subjects ≥9 years of age at screening (Visit 0) will complete the patient instrument: ItchRO(Pt). If a subject turns 9 years of age at any point after screening (Visit 0), they will not complete the ItchRO(Pt). Due to the expected lower number of subjects old enough to rate ItchRO(Pt), this measure will only be analyzed as an exploratory outcome.

Subjects and caregivers will be trained on the use of the electronic diary during the screening visit (Visit 0). Pruritus will be assessed and recorded twice daily, via ItchRO, beginning with the day after the screening visit (Visit 0) and every day throughout the duration of the study, as described in Appendix 2 and Appendix 3.

Exceptionally, a paper version may be made available in place of an electronic diary to accommodate cultural restrictions for a limited number of subjects/families to complete the required assessments. However, any requests for paper version will need to be proactively approved by the sponsor, who would provide any officially licensed instruments for data collection.

ItchRO(Obs)™

The severity of pruritus will be measured by the completion of the first question in the ItchRO(Obs) (how severe were your child's itch-related symptoms). The frequency of pruritus will be measured by completion of the third question in the ItchRO(Obs) (how much of the time was your child rubbing or scratching). Caregivers will rate the severity and frequency of pruritus using 5 choices to describe their itching condition. The sixth choice, "I don't know" will not count toward their severity or frequency score (categorized as missing data) and is included to account for rare occasions that the designated caregiver cannot observe the child, or the child could not communicate the severity or frequency of their itching condition. Capturing "I don't know" should be kept at an absolute minimum, as this will lead to missing data.

ItchRO(Pt)™

The severity of pruritus will be measured by the completion of the first question in the ItchRO(Pt) (how itchy did you feel). The frequency of pruritus will be measured by completion of the third question in the ItchRO(Pt) (how much of last night/today did feeling itchy make you rub or scratch). Subjects will rate the severity and frequency of pruritus using 5 choices to describe their itching condition.

During the 26-week treatment period, subjects/caregivers will be required to submit twice daily assessments using the electronic diary. The electronic diary will be returned to the study center at the final treatment evaluation visit (Visit 9).

7.2.2.2 Serum Bile Acids and Other Cholestasis Biomarkers

Blood samples will be collected as described in Table 1 to measure levels of cholestasis biomarkers including total sBA, sBA subspecies, C4, FGF-19, and autotaxin as well as liver-related parameters. Subjects are encouraged to fast at least 6 hours prior to collection (water intake is permitted if necessary but not recommended).

Total sBA and targeted bile acid subspecies will be quantified with liquid chromatography-mass spectrometry (LC-MS) methodology for exploratory assessments. In addition, screening total sBA will be assessed with an enzymatic assay for inclusion criteria evaluation. C4, a key intermediate in the pathway for bile acid synthesis from cholesterol, will be determined by a validated LC-MS/MS method.

Samples collected for biomarker evaluation may be utilized for additional assessments of related analytical methodology or other exploratory analysis.

7.2.3 Medical History, Disease History, and Medication History

At the screening visit (Visit 0) the investigator must record all clinically or medically relevant information, including PFIC disease history, regardless of how much time has elapsed since the date of any diagnosis.

Refer to Section 5.1 for full details on collection of medication history. Prior treatment information, including any prior treatments for PFIC, must be recorded on the appropriate CRF page.

7.2.4 Demographics

Subject demographic information including gender, age (full date of birth), and race will be collected during the screening visit (Visit 0). Variance in pediatric age-related reference ranges is significant, and the full date of birth is required to ensure correct interpretation of the laboratory report.

Local, country-specific guidelines should be followed where there is regional variation for collection of demographic information.

7.2.5 Safety

The name and address of each third-party vendor (e.g., clinical laboratory) used in this study will be maintained in the investigator's and sponsor's files.

7.2.5.1 Physical Examination

Abnormalities (including presence of hepatomegaly, splenomegaly, edema, ascites, jaundice, or scleral icterus) identified at the screening visit (Visit 0) will be documented in the subject's source documents. Significant changes after the screening visit (Visit 0) should be assessed as potential AEs. Refer to Section 8 for additional details.

Physical examination assessments at each visit should also include specific assessments for signs of hepatomegaly, splenomegaly, edema, ascites, jaundice, and scleral icterus.

7.2.5.2 Adverse Event Collection

At each study visit and subject contact, subjects or their caregivers will be questioned in a general way to ascertain if AEs have occurred since the previous visit (e.g., "Have you had any health problems since your last visit?"). Adverse events will be collected from the time informed consent is signed (refer to Section 8).

7.2.5.3 Vital Signs (including Height and Weight)

Vital signs include blood pressure, heart rate, temperature, and respiration rate. Blood pressure should be determined by cuff using the same method, the same arm, and the same position, following 5 minutes of rest throughout the study. Any deviations from baseline (Visit 1) vital signs that are deemed clinically significant in the opinion of the investigator are to be recorded as an AE.

Height will be collected at the screening visit (Visit 0) and at all site visits during the treatment period. Height will be measured by trained site staff in children who can stand on their own (generally ≥2 years) or not (usually <2 years) using a calibrated stadiometer or headboard, respectively, via 2 independent measurements, recorded to the nearest 0.1 cm (and a third measurement if values differ by >0.5 cm). The same stadiometer should be used for all study visit measurements.

Weight will be assessed, recorded to the nearest 0.1 kg, using a calibrated balance or electronic scale (children who can stand on their own, generally ≥2 years of age) or infant scale (children <2 years).

Standardized instructions based on regulatory guidance for height and weight measurements will be provided in the study manual.

7.2.5.4 Electrocardiograms

The following parameters will be collected: heart rate, RR interval, PR interval, QRS duration, uncorrected QT interval, rhythm (sinus or specify other rhythm), ECG normal or abnormal (if abnormal specify abnormality).

QT interval corrected using both Bazett's and Fridericia's formula (QTcB and QTcF) will be derived from data collected in the eCRF.

ECGs should be recorded in the supine position, if possible after at least 10 minutes rest to ensure a stable baseline. A standard 12-lead ECG will be performed at screening (Visit 0), baseline (Visit 1), at Visit 5 and Visit 9 (ET/EOT). ECGs will be recorded locally in source documents and results captured in the eCRF.

7.2.5.5 Clinical Laboratory Evaluations

All clinical laboratory evaluations will be performed according to the central laboratory manual. Reference ranges are to be supplied by the laboratory and will be used to assess the clinical laboratory data for clinical significance and out-of-range pathological changes. The investigator or sub-investigator must assess out-of-range clinical laboratory values for clinical significance, indicating if the value(s) is/are not clinically significant or clinically significant. Abnormal clinical laboratory values, which are unexpected or not explained by the subject's clinical condition, may, at the discretion of the investigator or sponsor, be repeated as soon as possible until confirmed, explained, or resolved.

Clinical laboratory assessments will be performed as listed in Table 4.

Unscheduled local laboratory values are not required to be collected in the CRF except for those that are directly relevant to the monitoring of an AE/SAE; collection of other values may be requested by the sponsor.

Table 4 List of Laboratory Analytes

Serum β-hCG (if indicated)	**Clinical Chemistry**	**Lipid Panel**[a]	**Urinalysis**
CBC with Differential	Albumin	Total cholesterol	pH
Hematocrit	ALP	LDL-C (direct)	Specific gravity
Hemoglobin	Amylase	HDL-C	Protein
MCV, MCH, MCHC	ALT (SGPT)	Triglycerides (TG)	Glucose
Red blood cells	AST (SGOT)		Ketones
Platelets	Bicarbonate	**Cholestasis Biomarkers**[a]	Bilirubin
White blood cells	Bilirubin, direct (conjugated)	Total serum bile acids (sBA)	Occult blood and cells
WBC Differential (% and absolute)	Total serum bilirubin (TSB)	Serum bile acid subspecies	Nitrite
Neutrophils	Blood urea nitrogen (BUN)	7α-hydroxy-4-cholesten-3-one (C4)	Urobilinogen
Eosinophils	Calcium	FGF-19	Leukocyte esterase
Basophils	Chloride	Autotaxin	Microscopic examination[b]
Lymphocytes	Creatinine		Oxalate
Monocytes	GGT	**Lipid Soluble Vitamins**[a]	Urinary creatinine
Coagulation	Glucose	25-hydroxy vitamin D	**Maralixibat Levels**
aPTT (sec)	Lipase	Retinol	Maralixibat in plasma
INR	Phosphate	Retinol binding protein (RBP)	**Marker of hepatocellular carcinoma**
PT (sec)	Potassium	α-Tocopherol	α-Fetoprotein (AFP)
	Sodium		
	Total protein		
	Uric Acid		
	Measured serum Osmolality		

AFP=α-fetoprotein; ALP=alkaline phosphatase; ALT=alanine aminotransferase; aPTT=activated partial thromboplastin time; AST=aspartate aminotransferase; β-hCG=beta human chorionic gonadotropin; FGF-19=fibroblast growth factor 19; FIB-4=fibrosis-4; GGT=gamma-glutamyl transferase; PT=prothrombin time; MCH=mean corpuscular hemoglobin; MCHC=mean corpuscular hemoglobin concentration; MCV=mean corpuscular volume; RBP=Retinol binding protein: SGOT=serum glutamic-oxaloacetic transaminase; SGPT=serum glutamic-pyruvic transaminase; WBC=white blood cell.

[a] Blood samples for the analysis of cholestasis biomarkers, lipid panel and lipid soluble vitamins should be drawn prior to administration of vitamin supplementation and as much as possible approximately 6 hours after food or formula (water intake is permitted if necessary but not recommended). Other biomarkers [e.g., lysophosphatidic acid (LPA),] may be measured. At the discretion of the sponsor, samples will be collected and appropriately stored for subsequent analysis, as needed.

[b] Will be performed on abnormal findings unless otherwise specified.

Serum bile acid results will remain blinded to sites and to the blinded study team until after database lock of study MRX-502 except for results at screening (Visit 0). Anion gap and osmolar gap as well as corrected sodium, α-Tocopherol/Total Lipids Ratio, Retinol/RBP Molar Ratio, FIB-4, and APRI will be calculated.

A serum storage sample will be collected as a back-up sample in case re-analysis is needed.

7.2.5.6 Pregnancy Test

A serum β-HCG pregnancy test is performed on all females of childbearing potential at the screening visit (Visit 0), and the final treatment evaluation visit (Visit 9) or ET visit. A urine pregnancy test is performed on all female subjects of childbearing potential at other visits as outlined in Table 1, or if pregnancy is suspected or on withdrawal of the subject from the study.

7.2.5.7 Ultrasound Liver Imaging

An abdominal ultrasound will be performed at the screening visit (Visit 0) and at Week 26 (Visit 9; EOT/ET) to determine the presence of any liver mass. A screening ultrasound is not required if the results of an ultrasound or liver MRI completed in the last 6 months are available, and the clinical status of the subject has not changed significantly since the time of the test.

7.2.6 Others

7.2.6.1 Health-related Quality of Life and Pruritus Assessment

For any health-related quality of life assessments with age-specific modules or instruments performed starting at screening (EDQ[Obs] or EDQ[Pt]), the subject's age at the screening visit (Visit 0) will be used for the determination of the appropriate module or instrument to be used for the study.

For any health-related quality of life assessments with age-specific modules or instruments performed starting at baseline (PedQL and PIS), the subject's age at the screening visit (Visit 0) will be used for the determination of the appropriate module to be used for the study.

The same module or instrument will be completed for the duration of the study, regardless of subsequent birthdays throughout the study.

7.2.6.1.1 Pediatric Quality of Life

The PedsQL™ is a questionnaire that will be administered to subjects and caregivers using the age-appropriate PedsQL module (refer to Appendix 4). Subjects aged 8 to 12 years, and 13 to 18 years will self-complete the PedsQL Child Report, and the PedsQL Teenager Report, respectively. Caregivers of subjects will complete the age-appropriate Parent PedsQL report (i.e., Report for Infants, Toddlers, Young Children, Children, and Teenagers).

The PedsQL is a validated, modular instrument designed to measure health-related quality of life (HRQoL) in children and adolescents (Varni et al., 2001). In addition to the core generic PedsQL module, the multidimensional fatigue and family impact questionnaires will also be administered as outlined in Table 1 using the age-appropriate module (refer to Appendix 4).

7.2.6.1.2 Caregiver Impression of Severity of Pruritus

The CIS is a questionnaire that will be administered to caregivers as outlined in Table 1 and Appendix 5. The CIS is designed to assess the caregiver's perception of itch severity of their children. The questionnaire will be administered with a recall period of 1 week.

7.2.6.1.3 Patient Impression of Severity of Pruritus

The PIS is a questionnaire that will be administered to subjects as outlined in Table 1 and Appendix 5. Subjects aged ≥9 years will self-complete the questionnaire. The PIS is designed to assess the subject's perception of their itch severity. The questionnaire will be administered with a recall period of 1 week.

7.2.6.1.4 Clinician Scratch Scale

The CSS provides an assessment of itch severity. The clinician's assessment of the subject's pruritus will focus on scratching and visible damage to the skin as a result of scratching as observed by the physician. The CSS uses a 5-point scale, in which 0 designates no evidence of scratching and 4 designates cutaneous mutilation with bleeding, hemorrhage and scarring. A clinician's assessment of pruritus made by the principal investigator or sub-investigator using the CSS will be recorded at screening, baseline (Visit 1), and at additional study visits as outlined in Table 1. To the extent possible, assessments should be made by the same individual during study visits.

7.2.6.1.5 Patient Impression of Change (PIC)

The PIC is designed to assess the subject's perception of his/her itching at Week 26 (EOT) (see Appendix 8) compared to his/her itching prior to the start of treatment with study drug. The PIC will be completed, by subjects who are 9 years of age or older at the Week 26 (EOT) visit (see Table 1).

7.2.6.1.6 Caregiver Impression of Change (CIC)

The CIC is designed to assess the caregiver's perception of the subject's itch related symptoms and xanthoma severity at Week 26 (EOT) compared to his/her itch related symptoms and xanthoma severity prior to the start of treatment with study drug (see Appendix 9). The CIC will be completed by all caregivers at the Week 26 (EOT) visit (see Table 1).

7.2.6.1.7 Exploratory Diary Questionnaire

Pruritus will be assessed using the EDQ caregiver/patient reported outcome measure administered as a twice daily electronic diary as described in Appendix 7. Caregivers for all subjects aged <9 years will complete the Observer instrument: EDQ(Obs). Subjects ≥9 years of age will complete the patient instrument: EDQ(Pt).

Subjects and caregivers will be trained on the use of the electronic diary during the screening visit (Visit 0). Pruritus will be assessed and recorded twice daily by subjects or caregivers,

via EDQ, beginning with the day after the screening visit and every day throughout the duration of the study, as described in Table 1.

During the 26-week treatment period, subjects/caregivers will be required to submit twice daily assessments using the electronic diary. The electronic diary will be returned to the study center at Visit 9 (EOT/ET).

7.2.6.2 Clinical Pharmacology Assessments

Blood samples (anticoagulant K3 EDTA) will be drawn pre-dose and 2.5 hours (with a 30-minute window) post-morning dose at Week 10 (Visit 5) and Week 26 (Visit 9; EOT/ET) to assess the plasma level of maralixibat. For Visit 5 and Visit 9, the subject dose should be taken while on-site, under supervision of the site staff.

Actual PK blood sample collection time versus time of dosing will be recorded. The investigator should ensure that every effort will be made to collect all PK blood samples at the precise protocol scheduled time. PK blood collection must not deviate from the nominal collection time set forth in the protocol by more than the allowable window (-30 minutes). Samples drawn outside the window will be considered a protocol deviation.

7.2.6.3 Healthcare Utilization

The number of hospitalizations related to underlying disease, length of stay for hospitalizations, and emergency room visits (days) and any surgeries and procedures specific to subject's PFIC condition will be collected during post-baseline visits as outlined in Table 1. The date of visit/admission, date of discharge, relationship/issue, and discharge status will be collected. The number of days the caregiver missed from work due to the events collected for healthcare utilization will also be collected.

7.2.7 Volume of Blood to Be Drawn from Each Subject

As detailed in Table 5, it is expected that approximately 104 mL of blood will be drawn during the study. According to the Ethical Considerations for Clinical Trials on Medicinal Products Conducted with the Paediatric Population, the volume of blood drawn from a subject should not exceed 3 % of the total blood volume during a period of 4 weeks and should not exceed 1% at any single time (Ethical Considerations, 2017). Should the volume of blood required for a single visit or a 30-day period exceed the maximum allowable amount, the investigator will draw blood in the priority order listed in Table 5 until the maximum amount has been reached. Lab draws missing due to the maximum allowable amount being reached will not be considered protocol deviations. Further instructions will be provided in the study manual or laboratory manual.

Table 5 Blood Sample Prioritization and Approximate Volume of Blood to Be Drawn from Each Subject

Order of Priority	Assessment		Minimal Sample Volume Required (mL)	Number of Samples	Total Volume (mL)
1	Total serum bile acid (tSBA)		1	9	9
2	Safety	Chemistry panel/β-hCG[a] (including lipid panel)	2.5	9	22.5
		Coagulation	1.8	9	16.2
		CBC with differential	1	9	9
		AFP[b]	0.7	3	2.1
		Lipid soluble vitamins	2.6	7	18.2
3	Pharmacokinetic samples		1	4	4
4	C4		1	5	5
5	Cholestasis biomarkers[c]		2	5	10
6	Serum storage sample		1	8	8
				Total mL	**104**

AFP=α-fetoprotein; β-hCG=beta-human chorionic gonadotropin; C4=7α-hydroxy-4-cholesten-3-one; eCRF=electronic case report form.
[a] β-hCG testing for females of child-bearing potential only.
[b] Results from local AFP drawn as standard of care may be entered in the eCRF instead of taking additional samples.
[c] Cholestasis biomarkers will include: autotaxin, FGF-19 and sBA subspecies.
Note: The amount of blood to be drawn for each assessment is an estimate. The amount of blood to be drawn may vary according to the instructions provided by the manufacturer or laboratory for an individual assessment; however, the total volume drawn over the course of the study should be approximately 104 mL.

8 ADVERSE AND SERIOUS ADVERSE EVENTS ASSESSMENT

8.1 Definition of Adverse Events, Period of Observation, Recording of Adverse Events

An AE is any untoward medical occurrence in a clinical investigation subject administered a pharmaceutical product and that does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product (ICH Guidance E2A 1995).

All AEs are collected from the time the informed consent is signed until the defined follow-up period stated in Section 7.1.4. This includes events occurring during the screening phase of the study, regardless of whether or not study medication is administered. For subjects who enroll in the extension study (MRX-503) prior to the 7-day safety follow-up visit (Week 27), AEs will be collected/reported in MRX-502 up to the time of consent to study MRX-503.

Where possible, a diagnosis rather than a list of symptoms should be recorded. If a diagnosis has not been made, then each symptom should be listed individually. All AEs should be captured on the appropriate AE pages in the CRF and in source documents.

All AEs must be followed to closure, regardless of whether the subject is still participating in the study. Closure indicates that an outcome is reached, stabilization achieved (the investigator does not expect any further improvement or worsening of the event), or the event is otherwise explained. When appropriate, medical tests and examinations are performed so that resolution of event(s) can be documented.

8.1.1 Treatment-Emergent Adverse Event (TEAE)

A treatment-emergent adverse event (TEAE) is defined as any event emerging or manifesting at or after the initiation of treatment with an IP or medicinal product or any existing event that worsens in either intensity or frequency following exposure to the IP or medicinal product.

8.1.2 Unexpected Adverse Event

An unexpected adverse event is an AE whose nature, severity, specificity, or outcome is not consistent with the term, representation, or description used in the Reference Safety Information (RSI). "Unexpected" also refers to the AEs that are mentioned in the IB as occurring with a class of drugs or as anticipated from the pharmacological properties of the product but are not specifically mentioned as occurring with the particular product under investigation.

The expectedness of AEs will be determined by the sponsor using the IB as the RSI. This determination will include considerations such as the number of AEs previously observed, but not on the basis of what might be anticipated form the pharmacological properties of a product.

8.1.3 Suspected Unexpected Serious Adverse Reaction (SUSAR)

Suspected Unexpected Serious Adverse Reaction (SUSAR) is defined as any suspected adverse reaction to study treatment (ie, including active comparators) that is both serious and unexpected.

The event(s) must meet all of the following:

- Suspected adverse reaction
- Serious
- Unexpected
- Assessed as related to study treatment

All SUSARs must be reported as described in Section 8.2.7.

8.1.4 Severity Categorization

The severity of AEs must be recorded during the course of the event including the start and stop dates for each change in severity. An event that changes in severity should be captured as a new event. Worsening of pre-treatment events, after initiation of study medication, must be recorded as new AEs (for example, if a subject experiences mild intermittent dyspepsia prior to dosing of study medication, but the dyspepsia becomes severe and more frequent after first dose of study medication has been administered, a new AE of severe dyspepsia [with the appropriate date of onset] is recorded on the appropriate CRF).

The medical assessment of severity is determined by using the following definitions:

Mild: A type of AE that is usually transient and may require only minimal treatment or therapeutic intervention. The event does not generally interfere with usual activities of daily living.

Moderate: A type of AE that is usually alleviated with specific therapeutic intervention. The event interferes with usual activities of daily living, causing discomfort but poses no significant or permanent risk of harm to the research subject.

Severe: A type of AE that interrupts usual activities of daily living, or significantly affects clinical status, or may require intensive therapeutic intervention.

8.1.5 Relationship Categorization

A physician/investigator must make the assessment of relationship to study medication for each AE. The investigator should decide whether, in his or her medical judgment, there is a reasonable possibility that the event may have been caused by the study medication. If there is no valid reason for suggesting a relationship, then the AE should be classified as "not related." Otherwise, if there is any valid reason, even if undetermined or untested, for suspecting a possible cause-and-effect relationship between the study medication and the occurrence of the AE, then the AE should be considered "related." The causality assessment must be documented in the source document.

The following additional guidance may be used to determine the relationship:

Term	Relationship Definition
Related	The temporal relationship between the event and the administration of the study medication is compelling and/or follows a known or suspected response pattern to that product, and the event cannot be explained by the subject's medical condition, other therapies, or accident.
Not Related	The event can be readily explained by other factors such as the subject's underlying medical condition, concomitant therapy, or accident and no plausible temporal or biologic relationship exists between the study medication and the event.

8.1.6 Outcome Categorization

The outcome of AEs must be recorded on the CRF. Outcomes are as follows:

- Fatal
- Not recovered/Not Resolved
- Recovered/Resolved
- Recovered/Resolved with Sequelae
- Recovering/Resolving
- Unknown

Action taken (i.e., dose increased, dose not changed, dose reduced, drug interrupted, drug withdrawn, not applicable, or unknown) will also be recorded on the AE CRF.

8.1.7 Symptoms of the Disease Under Study

Symptoms of the disease under study should not be classed as AEs as long as they are within the normal day-to-day fluctuation or expected progression of the disease and are part of the efficacy data to be collected in the study; however, disease progression resulting in study discontinuation should be recorded as an AE.

8.1.8 Clinical Laboratory and Other Safety Evaluations

A change in the value of a clinical laboratory, vital sign, or ECG assessment can represent an AE if the change is clinically relevant or if, during treatment with the study medication, a shift of a parameter is observed from a normal value to an abnormal value, or a further worsening of an already abnormal value. When evaluating such changes, the extent of deviation from the reference range, the duration until return to the reference range, either while continuing treatment or after the end of treatment with the study medication, and the range of variation of the respective parameter within its reference range, must be taken into consideration.

Laboratory abnormalities should not be classified as AEs as long as they are within the normal day-to-day fluctuation or expected progression of the disease and are part of the efficacy data to be collected in the study; however, significant worsening of laboratory parameters should be recorded as an AE. Laboratory abnormalities related to disease progression that result in interruption of study medication should be captured as AEs.

If, at the end of the treatment phase, there are abnormal clinical laboratory, vital sign, or ECG findings which were not present at the pretreatment assessment performed closest to the start of study treatment, further investigations should be performed until the values return to within the reference range or until a plausible explanation (e.g., concomitant disease) is found for the abnormal values.

The investigator should decide, based on the above criteria and the clinical condition of a subject, whether a change in a clinical laboratory, vital sign, or ECG parameter is clinically significant and therefore represents an AE.

8.1.9 Pregnancy

All pregnancies are to be reported from the time informed consent is signed until the defined follow-up period stated in Section 7.1.4.

Any report of pregnancy for any female study participant or the partner of a male subject must be reported within 24 hours to the sponsor or designated CRO medical monitor using the pregnancy report form.

The pregnant female study participant must be withdrawn from the study.

Every effort should be made to gather information regarding the pregnancy outcome and condition of the infant. It is the responsibility of the investigator to obtain this information within 30 calendar days after the initial notification and approximately 30 calendar days postpartum.

Pregnancy complications such as spontaneous abortion/miscarriage, elective abortion due to a congenital abnormality or a complication of pregnancy or a congenital abnormality diagnosed at any time during the pregnancy or post-partum are considered SAEs and must be reported using the serious adverse event form. An elective abortion of a normal pregnancy without complications is not considered an SAE.

In addition to the above, if the investigator determines that the pregnancy meets serious criteria, it must be reported as an SAE using the serious adverse event form as well as the pregnancy report form.

8.1.10 Abuse, Misuse, Overdose, and Medication Error

Abuse, misuse, overdose, or medication error (as defined below) must be reported to the medical monitor according to the SAE reporting procedure whether or not they result in an AE/SAE as described in Section 8.2. Note: The 24-hour reporting requirement for SAEs does not apply to reports of abuse, misuse, overdose, or medication errors unless these result in an SAE.

The categories below are not mutually exclusive; the event can meet more than 1 category.

- **Abuse** – Persistent or sporadic intentional intake of study medication when used for a non-medical purpose (e.g., to alter one's state of consciousness or get high) in a manner that may be detrimental to the individual and/or society
- **Misuse** – Intentional use of study medication other than as directed or indicated at any dose (Note: this includes a situation where the study medication is not used as directed at the dose prescribed by the protocol)

- **Overdose** – Intentional or unintentional intake of a dose of study medication exceeding the protocol-prescribed daily dose
- **Medication Error** – An error made in prescribing, dispensing, administration, and/or use of a study medication. For studies, medication errors are reportable to the sponsor only as defined below.

Cases of subjects missing doses of the study medication are not considered reportable as medication errors. The administration and/or use of the unassigned or expired study medication is/are always reportable as a medication error.

8.2 Serious Adverse Event Reporting Procedures

8.2.1 Reference Safety Information

The RSI for this study is the maralixibat chloride IB which the sponsor has provided under separate cover to all investigators.

8.2.2 Reporting Procedures

All SAEs must be reported by the investigator to the medical monitor within 24 hours of becoming aware of the event. Note: The 24-hour reporting requirement for SAEs does not apply to reports of abuse, misuse, overdose, or medication errors (see Section 8.1.10) unless they result in an SAE.

The investigator or designee must complete, sign, and date the serious adverse event form, verify the accuracy of the information recorded on the form with the corresponding source documents (Note: Source documents are not to be sent unless requested), and submit the form following the instructions in the study manual. Follow-up information should be reported as soon as available.

8.2.3 Serious Adverse Event Definition

A **serious adverse event** (SAE) is any untoward medical occurrence (whether considered to be related to study medication or not) that at any dose:

- Results in death
- Is life-threatening. Note: The term "life-threatening" in the definition of "serious" refers to an event in which the subject was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it was more severe.
- Requires inpatient hospitalization or prolongation of existing hospitalization. Note: Hospitalizations, which are the result of elective or previously scheduled surgery for preexisting conditions, which have not worsened after initiation of treatment, should not be classified as SAEs. For example, an admission for a previously scheduled

ventral hernia repair would not be classified as an SAE; however, complication(s) resulting from a hospitalization for an elective or previously scheduled surgery that meet(s) serious criteria must be reported as SAE(s).

- Results in persistent or significant disability/incapacity
- Is a congenital abnormality/birth defect
- Is an important medical event. Note: Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered an SAE when, based upon appropriate medical judgment, they may jeopardize the subject and may require medical or surgical intervention to prevent 1 of the outcomes listed in this definition. Examples of such medical events include allergic bronchospasm requiring intensive treatment in an emergency room or at home; blood dyscrasias or convulsions that do not result in inpatient hospitalization; or the development of drug dependency or drug abuse.

8.2.4 Serious Adverse Event Collection Time Frame

All SAEs (regardless of relationship to study) are collected from the time the subject signs the informed consent until the defined follow-up period stated in Section 7.1.4 and must be reported within 24 hours of the first awareness of the event. For subjects who enroll in the extension study (MRX-503) prior to the 7-day safety follow-up visit (Week 27), SAEs will be collected/reported in MRX-502 up to the time of the subject consent for MRX-503.

In addition, any SAE(s) considered "related" to the study medication and discovered by the investigator at any interval after the study has completed must be reported within 24 hours of the first awareness of the event.

8.2.5 Serious Adverse Event Onset and Resolution Dates

The onset date of the SAE is defined as the date the event meets serious criteria. The resolution date is the date the event no longer meets serious criteria, the date the symptoms resolve, or the date the event is considered chronic. In the case of hospitalizations, the hospital admission and discharge dates are considered the onset and resolution dates, respectively.

In addition, any signs or symptoms experienced by the subject after signing the informed consent form or leading up to the onset date of the SAE, or following the resolution date of the SAE, must be recorded as an AE, if appropriate.

8.2.6 Fatal Outcome

Any SAE that results in the subject's death (e.g., the SAE was noted as the primary cause of death) must be recorded on the CRF with an outcome of "fatal" and the date of death recorded as the resolution date. For all other events ongoing at the time of death that did not contribute to the subject's death, the outcome should be recorded as "not resolved," without a resolution date recorded.

For any SAE that results in the subject's death or any ongoing events at the time of death, unless another study medication action was previously taken (e.g., drug interrupted, reduced, withdrawn), the action taken with the study medication should be recorded as "dose not changed" or "not applicable" (if the subject never received study medication). The study medication action taken of "withdrawn" should not be selected solely as a result of the subject's death.

8.2.7 Regulatory Agency, Institutional Review Board, Ethics Committee, and Site Reporting

The sponsor is responsible for notifying the relevant regulatory authorities (US central IRBs/EU central ECs) of related, unexpected SAEs.

In addition, the sponsor is responsible for notifying active sites of all related, unexpected SAEs occurring during all interventional studies across the maralixibat program.

The investigator is responsible for notifying the local IRB, local EC, or the relevant local regulatory authority of all SAEs that occur, as required.

During the course of the study, the sponsor shall inform the regulatory authorities of any SUSARs occurring with maralixibat that occur globally as follows:

(a) if it is neither fatal nor life threatening, within 15 days after becoming aware of the information; and

(b) if it is fatal or life threatening, within 7 days after becoming aware of the information.

For fatal and life-threatening events, the sponsor shall, within 8 days after having informed the regulatory authorities, submit a complete report including an assessment of the importance and implication of any findings made.

In addition, local, country-specific guidelines should always be followed where there is regional variation in reporting.

8.3 Safety Monitoring Rules

8.3.1 General Guidelines

The following guidelines are provided for the monitoring of selected laboratory parameters.

Confirmation Guidance: At any time during the study, the initial clinical laboratory results meeting the safety monitoring criteria presented below must be confirmed by performing measurements on new specimens (retest).

Retests should be conducted by the central laboratory. Retests at a local laboratory are acceptable on an as-needed basis. Local laboratory results are sent to the investigator and medical monitor for review (these should include the normal range of the local laboratory). These results should be captured in the relevant CRF.

Retest collection should take place within 48 to 72 hours of the availability of the initial findings of potential concern. The results from the retest must be available prior to the next scheduled study visit or subject contact.

The baseline value for the purpose of the safety monitoring is the average between the value obtained during the screening visit (Visit 0) and the baseline visit (Visit 1), unless a confounder (e.g., intercurrent illness) causes variability outside of what is expected for the underlying disease based on the investigator's judgement (in which case a third value will need to be obtained).

8.3.2 Close Monitoring Criteria for Liver Parameters

Table 6 provides the criteria for close monitoring of liver parameters.

Table 6 Close Monitoring Criteria for Treatment-Emergent Elevated ALT and TSB

Parameter	Baseline Value	Close Monitoring Criteria
ALT	≤30 U/L	≥100 U/L
	>30 to ≤150 U/L	>3× BL or ≥250 U/L whichever comes first
	>150 to ≤450 U/L	≥(BL + 100 U/L)
TSB	Any	≥(BL + 3 mg/dL)

ALT=alanine aminotransferase; BL=baseline; TSB=total serum bilirubin.

Frequency of repeat measurements: If at any time during the study, an ALT or TSB result meets the close monitoring criteria, the measurement(s) must be re-confirmed. Subjects with a confirmed ALT or TSB level that is continuing to rise after meeting the close monitoring criteria should have their liver chemistry retested as clinically indicated, until levels stabilize or begin to recover.

Further investigation into liver chemistry elevations: For subjects with a confirmed increase in ALT or TSB level, meeting the close monitoring criteria, the following investigations should be considered, as clinically indicated:

- Close and frequent monitoring of liver enzyme and serum bilirubin tests as clinically indicated. Frequency of retesting can decrease if abnormalities stabilize, recover, or the study medication has been interrupted or discontinued and the subject has no symptomatic manifestations of hepatitis or immunological reaction.
- Symptoms as well as prior, current, and intercurrent diseases/illnesses

- Use of or recent change in use of concomitant treatment (including non-prescription medications, herbal and dietary supplement preparations), alcohol use, recreational drug use, and special diets.
- History for exposure to environmental chemical agents and travel
- Serology for viral hepatitis (HAV IgM, HBsAg, HCV antibody, HCVRNA, CMV IgM, and EBV antibody panel)
- Serology for autoimmune hepatitis (e.g., antinuclear antibody [ANA])
- AST, creatine phosphokinase (CPK), and lactate dehydrogenase (LDH)
- CBC with differential blood count (eosinophils)
- Reticulocyte count, PT/INR

Additional investigations of liver parameters, including gastroenterology/hepatology consultations, and/or hepatic imaging, may be performed at the discretion of the investigator, in consultation with the medical monitor

8.3.3 Guidelines for Interruption of Study Medication for Specific Liver Parameters

If the confirmed laboratory results meet any of the criteria below and the event is without an alternative explanation (i.e., intercurrent illness, disease progression, natural history of the disease), interruption of study medication should be considered:

The close monitoring criteria of both ALT and TSB are met simultaneously

OR

An increase of ALT or TSB is accompanied by signs and symptoms (e.g., nausea/vomiting, right upper quadrant pain) or an immunological reaction (e.g., rash, >5% eosinophilia, fever)

The following thresholds are met for either ALT or TSB:

Parameter	Baseline value	Observed change threshold
ALT	≤30 U/L	≥300 U/L
	>30 to ≤150 U/L	>(5× BL AND 300 U/L) OR ≥450 U/L whichever comes first
	>150 to ≤450 U/L	≥(BL + 300 U/L)
TSB	Any	≥(BL + 5 mg/dL)

ALT=alanine aminotransferase; BL=baseline; TSB=total serum bilirubin

Frequency of repeat measurements: Subjects with a confirmed ALT or TSB level meeting interruption criteria should have their liver chemistry tests retested every 72 hours or as clinically indicated, until levels stabilize or begin to recover.

Further investigation for increased liver chemistry results: For subjects with a confirmed increase in ALT or TSB level above the threshold, the investigations outlined in Section 8.3.2 should be considered, as clinically indicated.

Study medication may be re-started as long as the subject does not meet the dosing interruption criteria based on retest, does not meet criteria outlined in Section 8.3.4, liver parameters have returned to or near baseline, and an alternative explanation was found for the increased liver parameter (i.e., intercurrent illness, disease progression, or natural history of the underlying disease). The investigator and the medical monitor or qualified designee should confer as to whether additional close monitoring is needed.

8.3.4 Rules for Study Medication Discontinuation Following Abnormalities of Liver Parameters

Study medication should **not** be restarted:

- If the subject had a probable Drug-induced Livery Injury (DILI) in the opinion of the investigator and medical monitor; in case of disagreement, cases may be subject to independent expert adjudication
- If the subject reports signs/symptoms of acute hepatitis, liver decompensation (i.e., variceal hemorrhage, ascites, hepatic encephalopathy) or hypersensitivity reaction
- If the subject had liver parameters elevations meeting criteria in Section 8.3.3 without alternative explanations or that are inconsistent with the natural history of the disease
- After meeting stopping criteria subsequent to a re-challenge (i.e., meeting stopping criteria after re-introduction of study medication) that is likely related to study medication
- After 2 previous interruptions for safety reasons, unless there is a compelling reason why the subject should continue with dosing (i.e., interruptions were due to disease progression, natural history, intercurrent illness) and provided that the subject is still compliant with the study protocol as per Section 6.5

If study medication is permanently discontinued, the subject will be discontinued from the study and will be expected to complete the ET study procedures (Visit 9/ET).

8.3.5 Safety Monitoring for Lipid Soluble Vitamins

Lipid soluble vitamin (LSV) status will be assessed at every study visit per the schedule of assessments (Table 1); blood samples will be obtained at the study visit before the daily dose of vitamins is administered.

In subjects with LSV deficiency, study medication must be interrupted if clinical symptoms of deficiency occur. Clinical symptoms include, but are not limited to:

- Vitamin A: Night blindness, degeneration of the retina, xerophthalmia, follicular hyperkeratosis, keratomalacia, or Bitot's spots
- Vitamin D: Rickets, osteomalacia, costochondral beading, epiphyseal enlargement, cranial bossing, bowed legs
- Vitamin E: Numbness/tingling, decreased deep tendon reflexes, hemolytic anemia, limited extraocular movements, especially upward gaze

Subjects may restart dosing once clinical symptoms resolve and LSV levels begin to improve unless the subject has been previously discontinued for the same symptomatic LSV deficiency. Each subject may be re-challenged after symptomatic LSV deficiency only twice.

In the event of a confirmed level that falls either below the subject's baseline or below the lower limit of normal (refer to central laboratory manual), whichever is lower, for a vitamin (25-hydroxy vitamin D, retinol, retinol binding protein, α-tocopherol), or INR (as a proxy for vitamin K status), without clinical symptoms of deficiency, the investigator should make the appropriate modification to the subject's vitamin supplementation regimen (see also Section 8.3.6 for INR monitoring).

Appropriate starting doses and vitamins formulations should be used for new cases of LSV deficiency (see Table 7, adapted from Shneider et al., Pediatrics 2012). Doses should be increased, as clinically indicated, to optimize the LSV levels. Parenteral formulations should be used (where available) if enteral supplements are unsuccessful at improving LSV deficiency.

Table 7 Recommended Starting Doses for Lipid Soluble Vitamins

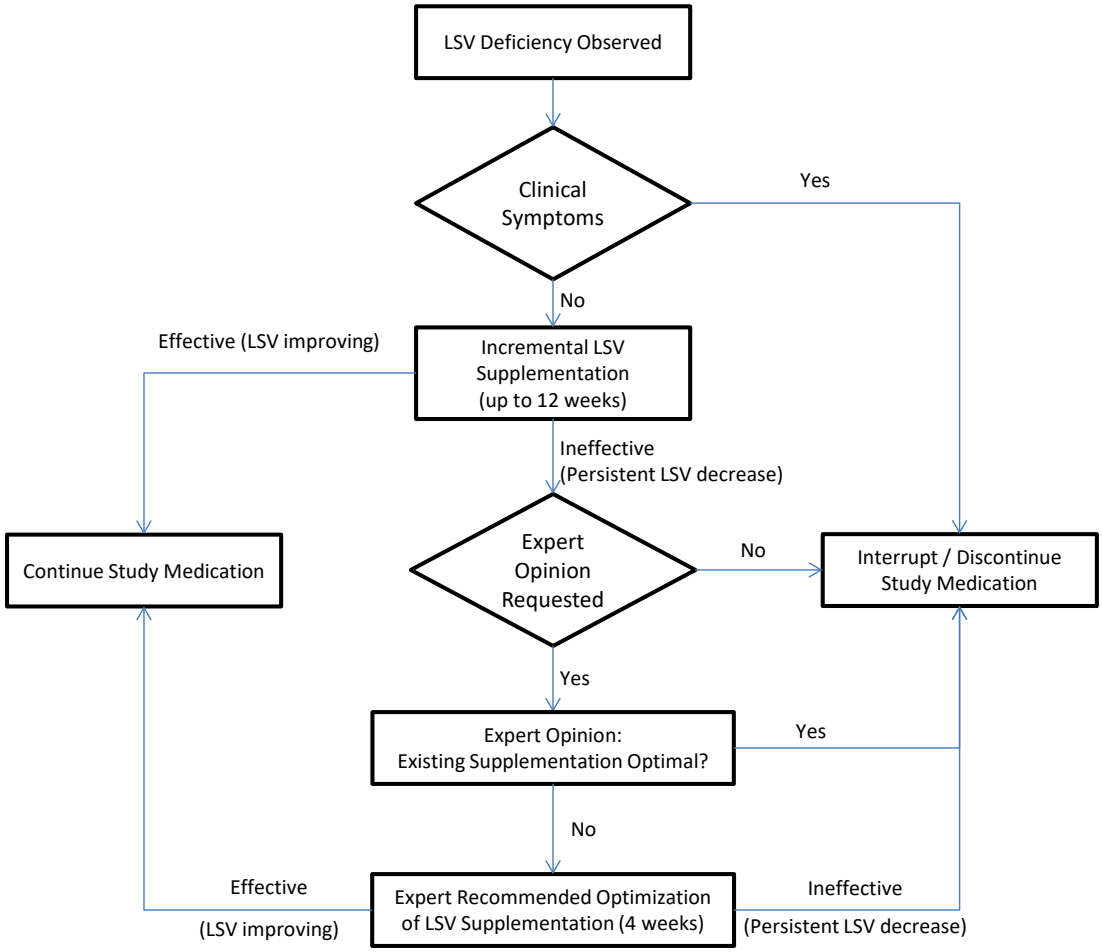
Vitamin	Starting Dose
Vitamin A	Starting dose 5000 IU orally of water miscible product (or appropriate intramuscular dose)
Vitamin D (25-hydroxy vitamin D)	Starting dose of 2000 IU orally daily of cholecalciferol (appropriate ergocalciferol or intramuscular dose may be used) or UVB phototherapy at the discretion of the investigator
Vitamin E (α tocopherol)	Starting dose of 25 IU/kg of TPGS (appropriate intramuscular dose vitamin E can be given if available)
Vitamin K	Starting dose of 2.5 mg orally for weight ≤30 kg or 5 mg for weight >30 kg (appropriate intramuscular or IV dose may be used)

The response to the change in regimen will be assessed at the next scheduled on-site study visit. LSV supplements are escalated in monthly intervals over a 3-month period. If during this period the LSV levels are persistently below the subject's baseline value or the lower limit of normal, whichever is lower, and there is an absence of clinical symptoms, the investigator will be given the option of:

- Consultation with a LSV expert to ensure the optimal doses/formulations of LSV supplements were administered
    - If, in the judgement of the expert, the doses/formulations of the administered supplements were optimal and yet their respective levels have not improved or further deteriorated, the subject will be discontinued from study medication.
    - If the expert recommends further optimization via supplementation, the investigator, in discussion with the medical monitor, may decide to continue study medication, administer the modified dose/formulation and re-check LSV levels at the next scheduled on-site visit. If, the levels of LSV levels have decreased further or have not improved, the subject will be permanently discontinued from study medication.
- Permanent withdrawal of the subject from the study.

Figure 2 provides a flow diagram representing the protocol recommendations for treatment of and discontinuation due to worsening of LSV deficiency.

Figure 2 Flow Diagram for Treatment of Worsening of Lipid Soluble Vitamin Deficiency



LSV=lipid soluble vitamins.

8.3.6 Safety Monitoring for Coagulation Panel Results

In the event of a confirmed laboratory result for INR >1.5 (BL ≤ULN) OR an increase of 0.4 over BL (BL >ULN) despite adequate Vitamin K supplementation, the investigator and the medical monitor may consider a temporary interruption of study medication. Dosing may resume when the INR falls below 1.5 or returns to the subject's baseline level.

8.3.7 Study Medication Discontinuation Rules for Diarrhea

Study medication should be discontinued if a subject has severe diarrhea that requires hospitalization and/or an IV or nutritional supplementation or leads to severe electrolyte disturbances.

9 DATA MANAGEMENT AND STATISTICAL METHODS

9.1 Data Collection

The investigators' authorized site personnel must enter the information required by the study CRF Completion Guidelines for all data requiring transcription of the source. A study monitor will visit each site in accordance with the monitoring plan and review the CRF data against the source data for completeness and accuracy. Discrepancies between source data and data entered on the CRF will be addressed by qualified site personnel. When a data discrepancy warrants correction, the correction will be made by authorized site personnel. Data collection procedures will be discussed with the site at the site initiation visit and/or at the investigator's meeting. Unscheduled assessments (e.g., local labs) are not required to be collected in the CRF except for those that are directly relevant to the monitoring of an AE/SAE; collection of other values may be requested by the sponsor.

The study will be monitored according to ICH GCP guidelines.

9.2 Clinical Data Management

A clinical database will be designed to collect data as specified in the data management plan. Quality control and data validation procedures are applied to ensure the validity and accuracy of the clinical database.

Data are to be reviewed and checked for omissions, errors, and values requiring further clarification using computerized and manual procedures. Data queries requiring clarification are to be communicated to the site for resolution. Only authorized personnel will make corrections to the clinical database, and all corrections are documented in an auditable manner.

9.3 Data Handling Considerations

Data that may potentially unblind the treatment assignment (i.e., maralixibat serum concentrations, treatment allocation, post-baseline sBA levels) will be handled with special care during the data cleaning and review process. Prior to unblinding, any data that may potentially unblind study site personnel or study team personnel will be presented as blinded information or otherwise will not be made available. If applicable, unblinded data may be made available to quality assurance representatives for the purposes of conducting independent audits.

9.4 Statistical Analysis Process

All pre-defined statistical analyses will be performed after the database is locked. All statistical analyses will be performed using SAS® (SAS Institute, Cary, NC, USA) version 9.4 or higher.

The SAP will provide the comprehensive statistical methods and definitions for the analysis of the efficacy and safety data, as well as describe the approaches to be taken for

summarizing other study information such as subject disposition, demographics and baseline characteristics, medical and PFIC disease history, study medication exposure, and prior medications. The SAP will also include define how missing data will be addressed. All study data will be presented in subject listings.

The SAP will be finalized prior to unblinding to preserve the integrity of the statistical analysis and study conclusions.

9.5 Validation Analysis Plan

A blinded analysis of the ItchRO(Obs) instrument will be performed before database lock according to the validation analysis plan (VAP). The blinded analysis will be used to estimate a threshold of clinically meaningful change (i.e., responder definition) in the 4-week average morning ItchRO(Obs) pruritus scores. The analysis will be performed by a group independent from both the study team and Mirum Pharmaceuticals. Blinded data will be used for this analysis; data will be collapsed across treatment groups.

Details of the analysis will be included in a standalone VAP, which will describe analyses that will be used to assess the measurement properties of the ItchRO(Obs) in alignment with scientific recommendations outlined in the FDA's Patient Focused Drug Development Guidance 3 (i.e., reliability, validity, ability to detect change, and anchor-based meaningful within-person clinically meaningful change criteria). Criteria for defining within-person clinically meaningful change for the ItchRO(Obs) will be detailed in the SAP.

9.6 Data Monitoring Committee

A data monitoring committee (DMC) will be involved in the management of this study. DMC meetings will be held periodically for the duration of the study. The purpose of the DMC is to review the progress of the study, with regard to safety and make recommendations to the sponsor to stop or modify the study if safety concerns are identified.

- The DMC will pay particular attention to the safety profile of the highest dose level (600 µg/kg BID). Should the DMC review determine that this dose is not safe, further dosing at this level, must be halted, and the relevant regulatory authority informed.
- The DMC will also be able to request the frequency of safety laboratory tests be increased during the dose escalation period if this is considered necessary based on their review of safety data.

In addition to scheduled meetings, the DMC will convene for ad hoc meetings in case of any safety concerns arising during the conduct of the study. In particular, an ad hoc DMC meeting will be held within 2 weeks if ≥2 similar possibly/probably related SAEs or Grade 4 TEAEs are observed. The DMC will perform an unblinded review of these cases and provide recommendations to the sponsor regarding the continuation of specific dose levels, or the study overall. Study drug would be reduced or stopped in those subjects who

experienced these triggering events, as per investigator judgment; study medication in all other participants will continue unless and investigators are otherwise advised by the DMC.

Further details regarding the DMC can be found in the DMC charter, which will be available prior to the enrollment of the first subject. There is no study stop planned based on the efficacy results during the study and no alpha spending will occur prior to the completion of the study.

9.7 Sample Size Calculation and Power Considerations

The sample size was calculated to ensure enrollment of an adequate number of PFIC2 participants who fulfill criteria for the primary cohort (primary endpoint). The sample size is chosen to provide at least 80% of power for the analysis of the primary endpoint.

The primary efficacy endpoint is the mean change from baseline in the average morning ItchRO(Obs) severity score for the last 12 weeks of the study (Weeks 15–26).

The following assumptions were made in estimating the sample size for study MRX-502:

- Between-treatment group (placebo to active) LS mean difference of 0.663
- Pooled SD of 0.563
- Effect size of 1.178 in the mean change from baseline in the 4-week average morning ItchRO(Obs) score

The estimations of LS mean difference, SD and effect size for the sample size justification are based on ItchRO(Obs) data from the maralixibat Phase 2 studies. Maralixibat-treated subjects with PFIC2 from Study LUM001-501 were compared to placebo-treated subjects with ALGS from Studies LUM001-301 and LUM001-302 as no placebo data are available from Study LUM001-501. Since impact and symptoms of itching reported by PFIC patients are also reported by ALGS patients, (i.e., mood changes, scratching, difficulty falling asleep, skin damage, blood from scratching etc.), it is reasonable to assume that the ItchRO outcomes from the ALGS placebo-controlled studies will inform outcomes for this PFIC study due to the similarities in the features of pruritus. Two fundamental assumptions were made: A) the placebo effect in subjects with ALGS is similar to the one expected in subjects with PFIC2. Subjects with ALGS in study LUM001-301/302 had a slightly higher BL ItchRO(Obs) value compared to subjects with PFIC in Study LUM001-501. As subjects with higher BL pruritus generally report a higher placebo response, the use of ALGS subjects from study LUM001-301/302 would therefore lead to a slightly more conservative assumption of the expected effect and sample size. The 2 populations were similar in the other BL characteristics. Besides from the small difference in BL characteristics there are no obvious reasons to believe why the placebo response in subjects with ALGS and PFIC should be difference. Therefore, the comparison of ALGS placebo subjects with treated PFIC subjects is considered a valid approximation for the purpose of assessing the effect size, data variability and the calculation of the sample size. B) the placebo effect does not change after

3 months of treatment; as Studies LUM001-301 and -302 were 13-week studies, data beyond 13 weeks were imputed using the last observation carried forward.

With the assumed values, a total of 26 complete subjects (13 subjects in each treatment group) will be required to provide 80% power for the comparison of the primary endpoint measure between the maralixibat treatment group and placebo, based on a 2-sided, 2-sample t-test at the 0.05 level of significance.

Including a 10% dropout rate based on the previous PFIC study and rounding up to the next even number, approximately 30 participants with PFIC2 (15 subjects in each treatment group) will be randomized for the primary analysis.

9.8 Analysis Populations

The following efficacy analysis populations are defined:

The **intent-to-treat (ITT) population** will consist of all randomized subjects.

The **per-protocol population** will consist of all subjects in the ITT population who receive at least 1 dose of study medication and do not have any major protocol violations or deviations. Major protocol violations/deviations will be identified prior to database lock.

The primary and secondary efficacy endpoints will be analyzed for each of the analysis populations described above. The primary analysis on the primary efficacy endpoint will be conducted in participants with PFIC2 in the ITT population. The primary analysis on the secondary efficacy endpoints will be conducted in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) in the ITT population.

Any exploratory efficacy endpoints will be analyzed in the ITT population. Exploratory analyses may be performed on participants with PFIC2, PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6), and other PFIC populations of interest. The final list of exploratory endpoints will be described in the SAP.

The **safety population** will consist of all subjects who receive at least 1 dose of study medication.

Safety, tolerability, and PK analyses will be conducted in the safety population.

Because PFIC is a rare disease, siblings are allowed to enroll in the study. Each sibling (if any) will be considered for each analysis population described above. A sensitivity analysis on the primary and secondary efficacy endpoints will include only 1 sibling in each efficacy analysis population. The choice of sibling will be made at random, and the method will be pre-specified in the SAP.

9.8.1 PFIC Populations for Analyses

The following definitions will be used for the PFIC populations in the analyses of the various endpoints in this study:

Primary and Secondary Endpoints:

- **PFIC2**: Participants with PFIC2 who fulfill criteria for the primary cohort (see Section 4)
- **PFIC**: Participants with PFIC1, PFIC2 (who fulfill criteria for the primary cohort), PFIC3, PFIC4, PFIC5, and PFIC6.

Exploratory Endpoints:

PFIC2 and PFIC populations (as above) and

- **PFIC1**: Participants with PFIC1
- **PFIC3**: Participants with PFIC3
- **PFIC4:** Participants with PFIC4
- **PFIC5:** Participants with PFIC5
- **PFIC6:** Participants with PFIC6
- **Truncated PFIC2:** Participants with PFIC2 with mutations leading to premature protein truncation or failure of protein production
- **All-participants**: All participants enrolled

Each of the above-defined PFIC populations, with the exception of the All participants, will exclude any participants with biliary surgery at baseline and genotype variations affecting only 1 allele (heterozygous).

In the PFIC population (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6), participants with biallelic disease causing variation in ATP8B1 (PFIC1), ABCB11 (PFIC2), ABCB4 (PFIC3), TJP2 (PFIC4), NR1H4 (PFIC5), and MYO5B (PFIC6) may be included. The number of participants with each variation will depend on the genotype prevalence within the enrolled study population.

9.9 Subject Disposition, Demographics, and Baseline Characteristics

Subject disposition, demographics and baseline characteristics, PFIC disease history, prior medications, and treatment exposure and compliance will be summarized by treatment group.

Study completion status and reasons for discontinuation will also be displayed. The analysis population(s) to be used for these summaries will be provided in the SAP.

Medical and surgical history and protocol violations/deviations will be presented in subject listings.

9.10 Efficacy Analyses

The primary and secondary efficacy analysis will be performed for the ITT population. Analysis of the primary and secondary efficacy endpoints will also be performed in the per protocol population. Analyses on the exploratory efficacy endpoints will be performed in the ITT population. Exploratory analyses may be performed in participants with PFIC2, PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6), and other PFIC populations of interest. Final list of exploratory endpoints will be described in the SAP.

For the PFIC population (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6), PFIC type will also be used as a covariate in the analysis. For this analysis, due to the small sample sizes, PFIC4, PFIC5, and PFIC6 will be grouped together.

In general, the baseline is defined as all data collected at Visit 1 (Day 0) or the last assessment before the first dose of the study medication if the data are not collected at Visit 1 (Day 0). Baseline ItchRO and EDQ average scores are defined as the 4-week average score prior to the first dose of study medication.

For all efficacy analyses, subjects will be analyzed by the randomized treatment group assignment (maralixibat or placebo), based on the ITT principle that asserts that the effect of a treatment can be best assessed by evaluating on the basis of the intention to treat a subject (i.e., the planned treatment regimen) rather than the actual treatment given. All efficacy data will be presented in subject listings.

9.10.1 Primary Efficacy Endpoint

The primary efficacy endpoint is defined as the mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26, using 4-week average morning ItchRO(Obs) severity scores.

Weekly scores will be calculated relative to the initiation of treatment. For each week, a weekly score will be calculated if at least 4 days' worth of data are provided. If more than 50% of scores are missing for a particular week, the weekly average score will be treated as missing. Average scores will be calculated for the following periods: Through Week 6 (42 days), Weeks 7–10, 11–14, 15–18, 19–22, and 23–26 (28 days). For each period, the average of the weekly scores will be calculated. At least 50% of the weekly scores must be available in order to compute an average for the period. The baseline average morning ItchRO(Obs) score is defined as the 4-week average morning ItchRO(Obs) severity score prior to the first dose of study medication.

Primary Analysis for Primary Efficacy Endpoint

For this study, the primary estimand is the improvement in pruritus measured as change from baseline in the average morning ItchRO(Obs) severity score in the maralixibat treatment group relative to the placebo group. In the course of the 26-week randomized treatment period, participants may be exposed to possible known or unknown intercurrent events that could possibly impact the interpretation of the measures associated with the clinical question of interest, such as treatment discontinuation due to a specific adverse effect or perhaps a lack of effect. A treatment policy estimand will be utilized for the analysis where all data will be incorporated. Intercurrent events will be handled as follows:

- Change in or administration of any allowed concomitant medication will be handled using a treatment policy strategy.
- Administration of any prohibited concomitant medication will be handled using a hypothetical strategy assuming that the prohibited medications are not available and hence were not taken, so as to assess the treatment effect for the randomized treatment groups in addition to allowed concomitant medication.
- Discontinuation of randomized treatment due to all of the following reasons will be handled using a treatment policy strategy:
    - Lack of efficacy
    - Safety concerns
    - Lack of compliance with study treatment
- Discontinuation of randomized treatment due to other reasons (including loss to follow-up) will be handled using a hypothetical strategy, assuming that randomized treatment had not been discontinued.
- Liver transplant or death: Neither of these are expected in the study. In the unlikely event that either of these events occur, they will be analyzed using a composite strategy, assuming worst possible score for all subsequent assessments.

The primary analysis on the primary efficacy endpoint will be conducted in participants with PFIC2 who fulfill criteria for the primary cohort in the ITT population. A restricted maximum likelihood (REML)-based mixed-effects model for repeated measures (MMRM) will be used as the primary analysis method. The repeated measures include post-baseline visits during the dose escalation phase (i.e., Week 6) and stable dosing phase (i.e., Weeks 10, 14, 18, 22, and 26), with change from baseline in the 6- or 4-week average morning ItchRO(Obs) severity score as the dependent variable. The MMRM model will include the fixed, categorical effects of treatment group, visit, and treatment group-by-visit interaction as well as the continuous, fixed covariates of baseline 4-week average morning ItchRO(Obs) severity score and the baseline score-by-visit interaction.

The unstructured variance/covariance matrix will be used to model the variances and covariances for the 6 time points (periods) included in the model. The unstructured variance/covariance does not impose any restrictions on the pattern of the matrix elements. Every attempt (e.g., relaxing the convergence criteria, increasing the iteration limit, choosing reasonable starting values for the estimates) will be made to ensure convergence using the unstructured modeling of within-subject correlations. However, if the numerical algorithm for estimation of the mixed model fails to converge, the following variance/covariance matrix structures will be used in the following order: 1) heterogeneous Toeplitz, 2) heterogeneous autoregressive of order 1, and 3) heterogeneous compound symmetry. The first (co)variance structure which does not have a convergence problem will be the one used for the analysis. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom (Kenward and Roger, 1997).

The primary efficacy analysis will compare maralixibat and placebo using the contrast (difference in least squares [LS] means) between treatment groups across the last 12 weeks of the study (i.e., Weeks 15–18, 19–22, and 23–26 combined). The analytical solution of the overall treatment effect obtained from MMRM is an equally weighted average of the 3 individual period-specific estimates over the time period of interest (i.e., the last 12 weeks of the study). Significance tests will be based on LS means using a 2-sided significance level (2-sided 95% confidence intervals [CIs]).

The null hypothesis for the primary efficacy endpoint of the equality of maralixibat and placebo is:

H_{01}: mean change in average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26 in the 2 treatment groups are equal

The null hypothesis of equal treatment effect will be rejected if the statistical analysis results in a 2-sided p-value for treatment over the last 12 weeks of the study is ≤ 0.05. Least squares (LS) means will be calculated for each treatment group for each post-baseline period in the model. The difference between maralixibat and placebo change from baseline in average morning ItchRO(Obs) severity score will be estimated, with the corresponding 2-sided 95% CI constructed for each period, and the last 12 weeks (i.e., last 3 periods) of the study combined. The change from baseline LS means with standard error, 95% CI for the LS means, p-value for testing if the LS mean is zero, LS mean difference between treatment groups (maralixibat minus placebo) with standard error, 95% CI for the LS mean difference, and p-value for testing if the treatment LS means are equal will be presented.

The trial will be claimed successful if the hypothesis of no treatment effect on the primary efficacy endpoint in the ITT population is rejected at the 0.05 (2-sided) significance level.

Rationale for Primary Analysis Method

The MMRM method has been demonstrated extensively as an appropriate choice for the primary analysis in longitudinal confirmatory clinical trials with continuous endpoints (Mallinckrodt et. al., 2008). This analysis method, which is from a broader class of direct-likelihood analyses methods, makes use of fully and partially observed data sequences

from individual subjects by estimating the covariance between data from different time points (Molenberghs and Kenward, 2007). Further, it is often useful to implement MMRM using an unstructured approach to modeling both the treatment-by-time means and the (co)variances, leading to what is essentially a multivariate normal model wherein treatment group means at the primary time point are adjusted to reflect both the actual observed data and the projected outcomes from the subjects with missing data (Cnaan et al., 1997, Molenberghs et al., 2004, Molenberghs and Kenward, 2007).

Despite careful planning and study conduct, the occurrence of missing data cannot be completely eliminated. As a direct likelihood method, the MMRM method is a preferred approach for handling missing data in such designs. MMRM is a full multivariate model in nature, which avoids potential bias as a predetermined model and operates in a more general missing at random (MAR) framework (Mallinckrodt et al., 2001). Data are considered MAR if, conditional upon the independent variables in the analytic model, the missingness depends on the observed outcomes of the variable being analyzed but does not depend on the unobserved outcomes of the variable being analyzed. This assumption implies that the behavior of the post dropout observations can be predicted from the observed variables, and therefore that treatment effect can be estimated without bias using the observed data (European Medicines Agency, 2010). For studies of missing data in a controlled clinical trial setting, MAR is usually considered as a plausible underlying missing mechanism (Molenberghs and Kenward, 2007; Siddiqui et al., 2009, Mallinckrodt et al., 2008, Mallinckrodt et al., 2013). The assumption of MAR is often reasonable because, particularly in longitudinal studies wherein the evolution of treatment effects is assessed by design over time, the observed data and the models used to analyze them can explain much of the missingness (Little and Rubin, 1987, Verbeke and Molenberghs, 2000). This point may be especially relevant in well-controlled studies, in which extensive efforts are made to observe all outcomes and factors that influence them while subjects are following protocol-defined procedures. Thus, longitudinal clinical trials by their very design aim to reduce the amount of missing not at random (MNAR) data (missingness explained by unobserved responses), thereby increasing the plausibility of MAR (Mallinckrodt et al., 2008).

Sensitivity and Supportive Analysis for Primary Efficacy Endpoint

Sensitivity and supportive analysis will be performed on the primary efficacy endpoint to quantify the possible impact of missing data and the use of potentially correlated sibling data (if needed), and to demonstrate the robustness of the conclusions. As a sensitivity analysis, the primary efficacy endpoint will be analyzed in PFIC2 participants who fulfill criteria for the primary cohort in the ITT population using the MMRM model, described above.

Summary statistics on average morning ItchRO(Obs) severity scores by treatment group and visit, including the end of study visit (Week 26/EOT), will also be presented. Additionally, a figure with the LS mean ±SE of change from baseline in the average morning ItchRO(Obs) severity score will be presented by treatment group and visit.

Missing Data

Although the assumption of MAR, as used for the primary efficacy analysis method, is often reasonable in clinical trials, the possibility of MNAR data cannot be ruled out. Therefore, analysis valid under MNAR is desired. Both a MNAR- and MAR-based analyses will be the basis upon which sensitivity of missing data are assessed.

Sensitivity analyses to deal with missing data will use multiple imputation (MI) methods where missing values are imputed individually under both a plausible MNAR and MAR scenario. Multiple imputation is a simulation-based approach where missing values are replaced using an appropriate stochastic model given the observed data and covariates, creating multiple completed data sets. These completed datasets are then analyzed using standard analysis methods (i.e., MMRM), and the different parameter estimates across the datasets are then combined to produce a unique point estimate and standard error taking into account the uncertainty of the imputation process. The following 2 sensitivity analysis models, one based on the standard MAR approach and the other based on the MNAR approach, will be used to examine robustness of the primary analysis results.

Complete details of the multiple imputation procedures, including the methodology, number of imputations, and the seed of the pseudorandom number generator used to randomly generate imputations for the missing values, will be pre-specified in the SAP prior to unblinding.

Standard MAR Imputation

The MAR imputation model will impute missing values using a regression-based multiple imputation model (Little and Yau, 1996). For subjects with complete data up to a particular visit, a multiple regression model will be fit that includes the outcome at that visit as the dependent variable and observed data (e.g., outcomes at previous visits, treatment, and baseline) as independent variables. Separate models will be similarly constructed for each visit. Using these regression models, a missing value for a subject at a particular visit will be imputed as a draw from the predictive distribution given the outcomes at previous visits (some possibly imputed), treatment group, etc. This process will be repeated a given number of times (as specified in the SAP), resulting in the same number of complete analysis data sets. The MMRM analyses will be performed separately for each of the completed analysis data sets, and the results will be combined into one multiple imputation inference (Little and Yau, 1996, Schafer, 1997). This strategy is appropriate for data sets that have a monotone missing data pattern. If the data set does not precisely have this pattern, a monotone data augmentation method, as described in the SAP, will be used to impute the small amount of missing data that is required to make the missing data pattern monotone before applying the MI algorithm described above.

Placebo-Based Imputation

The placebo-based imputation analysis model assumes a MNAR mechanism and is within the pattern-mixture model framework. The pattern-mixture model approach models the distribution of a response as the mixture of a distribution of the observed responses and a

distribution of the missing responses. Conceptually, pattern mixture models typically assess the outcome variable separately for different groups (patterns) and then combine results across groups for final inference (Mallinckrodt et al., 2008, Ratitch and O'Kelly, 2011). The pattern-mixture model approach will be applied in the SAS MI procedure by using the MNAR statement. The MNAR imputation model will specify a subset of observations to derive the imputation model, which is used for pattern imputation based on the placebo group.

The underlying assumption is that a subject on the active treatment with missing data follows the distribution of the placebo responses, i.e., the means and the intra-subject correlations based on the placebo responses will apply. This sensitivity analysis stress-tests the MAR assumption that withdrawals will tend to have efficacy similar to subjects who remain in the study in their respective treatment arm. On the contrary, the placebo-based pattern imputation assumes MNAR; that after discontinuation, subjects on the IP will tend to have efficacy close to subjects on placebo.

The imputations for this model are based on the distribution of placebo group responses over time. That is, an imputation model for the missing observations in the maralixibat treatment group is constructed not from the observed data in the active treatment group but rather from the observed data in the placebo treatment group. This model is also the imputation model that is used to impute missing observations in the placebo group.

Sibling Data

If siblings are randomized and evaluable for efficacy, the primary analysis method described above for the primary efficacy endpoint will be repeated as a sensitivity analysis in which only one of sibling pairs or multiples will be used. The choice of which sibling to use in the sensitivity analysis will made before the MRX-502 study is unblinded. The choice of sibling will be made at random.

9.10.2 Adjustment for Multiplicity

In order to maintain study-wide Type I error control, a hierarchical testing procedure will be used in the comparisons between maralixibat and placebo on the primary and secondary efficacy endpoints in the ITT population. The hierarchical order for testing the null hypotheses is as follows:

1. Primary: Mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26 in participants with PFIC2
2. Secondary: Mean change in total sBA level between baseline and average of Weeks 18, 22, and 26 in participants with PFIC2
3. Secondary: Mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26 in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)

4. Secondary: Mean change in total sBA level between baseline and average of Weeks 18, 22, and 26 in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
5. Secondary: Proportion of ItchRO(Obs) responders from Week 15 to Week 26 in participants with PFIC2 using the average value from the three 4-week periods (Weeks 15-18, 19-22, and 23-26)
6. Secondary: Proportion of sBA responders from Week 18 to Week 26 in participants with PFIC2, using the average value from Weeks 18, 22, and 26 values
7. Secondary: Proportion of ItchRO(Obs) responders from Week 15 to Week 26 in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)
8. Secondary: Proportion of sBA responders from Week 18 to Week 26 in participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6)

Because the testing is sequential, the type I error rate of α=0.05 is maintained. Failure at any stage in the sequence implies no type I error control for the additional subsequent tests. All p–values for each comparison without adjustments will be provided in the summary tables for informational purposes.

The effect of such a procedure is that no confirmatory claims can be based on the endpoint(s) that have a rank lower than that endpoint whose null hypothesis was the first that could not be rejected. Those tests will be considered exploratory and no formal conclusions will be drawn.

9.10.3 Secondary Efficacy Endpoints

The secondary efficacy endpoints are defined as:

- Mean change in total sBA level between baseline and average of Weeks 18, 22, and 26
- Mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26 in participants with PFIC
- Proportion of ItchRO(Obs) responders from Week 15 to Week 26
  - For the purpose of determining response, the average value from the three 4-week periods (Weeks 15–18, 19–22, and 23–26) will be computed and used.

    <u>ItchRO(Obs) responder definition</u>:

    Criteria for clinically meaningful change in the ItchRO(Obs) (responder definition) will be established by the VAP and described in detail in the SAP.
- Proportion of sBA responders from Week 18 to Week 26
  - For the purpose of determining response, the average value from Weeks 18, 22, and 26 values will be computed and used.

sBA responder definition:

- ≥75% reduction in sBA

OR

- sBA level <102 µmol/L (applies only if baseline sBA level was >102 µmol/L)

Serum bile acid from each scheduled clinic visit will be used in the analysis of mean change from baseline in sBA.

Analysis for Secondary Efficacy Endpoints

Secondary efficacy endpoints will be analyzed in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) as described in Section 9.10. Analysis similar to that described for the primary efficacy endpoint, including the sensitivity analyses, will be performed for each of the secondary efficacy endpoints.

For the binary (responder) outcomes, the number and proportion of subjects that are considered a "responder" will be summarized by treatment group at the end-of-study visit (Week 26/EOT). Barnard's exact test will be used to calculate the p-value for the difference between treatment groups.

The null hypotheses for the secondary efficacy endpoints of the equality of maralixibat and placebo are:

H_{02}: mean change in total sBA level between baseline and average of Weeks 18, 22, and 26 on participants with PFIC2 in the 2 treatment groups are equal

H_{03}: mean change in the average morning ItchRO(Obs) severity score between baseline and Week 15 through Week 26 on participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) in the 2 treatment groups are equal

H_{04}: mean change in total sBA level between baseline and average of Weeks 18, 22, and 26 on participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) in the 2 treatment groups are equal

H_{05}: proportion of ItchRO(Obs) responders from Week 15 to Week 26 on participants with PFIC2 in the 2 treatment groups are equal using the average value from the three 4-week periods (Weeks 15-18, 19-22, and 23-26)

H_{06}: proportion of sBA responders from Week 15 to Week 16 on participants with PFIC2 in the 2 treatment groups are equal, using the average from Weeks 18, 22, and 26 sBA values

H_{07}: proportion of ItchRO(Obs) responders from Week 15 to Week 26 on participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) in the 2 treatment groups are equal

H_{08}: proportion of sBA responders from Week 18 to Week 26 on participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6) in the 2 treatment groups are equal

A hierarchical testing procedure, as described in Section 9.10.2, will be used in the comparisons between maralixibat and placebo on the primary and secondary efficacy endpoints in the ITT population using the primary analysis method (as described for the primary endpoint).

All tests will be performed as 2-sided tests at the 0.05 level of significance.

9.10.4 Exploratory Efficacy Endpoints

The final exploratory endpoint list will be described in detail in the SAP. Analyses on the exploratory efficacy endpoints will be performed in the ITT population. Exploratory analyses may be performed in participants with PFIC2, PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5, and PFIC6), and other PFIC populations of interest as described in the SAP.

- Mean change in liver-associated laboratory test level between baseline and Week 15 through Week 26
- Mean change from baseline in liver-associated laboratory tests to Weeks 2, 6, 10, 14, 18, 22, and 26
- Proportion of participants with elevated liver biochemistry (i.e., >ULN at baseline), whose liver biochemistry normalize (≤ULN) at Weeks 2, 6, 10, 14, 18, 22, and 26
- Mean change from baseline in FIB-4, APRI, MELD and PELD score to Weeks 2, 6, 10, 14, 18, 22, and 26
- Proportion of days with a morning, evening, and highest daily ItchRO(Obs) severity score ≤1 at the participant level from baseline to Week 6 (overall and weekly), Weeks 7 10, 11-14, 15-18, 19-22, 23-26, and 15-26
- Proportion of days with a morning, evening, and highest daily ItchRO(Obs) frequency score ≤1 at the participant level from baseline to Week 6 (overall and weekly), Weeks 7-10, 11-14, 15-18, 19-22, 23-26, and 15-26
- Proportion of days with a morning, evening, and highest daily ItchRO(Pt) severity score ≤1 at the participant level from baseline to Week 6 (overall and weekly), Weeks 7-10, 11-14, 15-18, 19-22, 23-26, and 15-26

- Proportion of days with a morning, evening, and highest daily ItchRO(Pt) frequency score ≤1 at the participant level from baseline to Week 6 (overall and weekly), Weeks 7-10, 11-14, 15-18, 19-22, 23-26, and 15-26
- Mean change in the average morning ItchRO(Obs) frequency score between baseline and Week 15 through Week 26
- Mean change from baseline in average (6-week average for Week 6 and 4-week average afterwards) morning, evening, and highest daily ItchRO(Obs) severity and frequency scores to Weeks 6, 10, 14, 18, 22, 26, and Weeks 15-26 combined
- Mean change from baseline in weekly average morning, evening, and highest daily ItchRO(Obs) severity and frequency scores to each post-baseline week, and Weeks 15-26 combined
- Mean change from baseline in average (6-week average for Week 6 and 4-week average afterwards) morning, evening, and highest daily ItchRO(Pt) severity and frequency scores to Weeks 6, 10, 14, 18, 22, 26, and Weeks 15-26 combined
- Mean change from baseline in weekly average morning, evening, and highest daily ItchRO(Pt) severity and frequency scores to each post-baseline week, and Weeks 15-26 combined
- Mean change from baseline in CSS, CIS, and PIS at Weeks 2, 4, 6, 10, 14, 18, 22, and 26
- Proportion of participants in each change from baseline category of the CSS, CIS, and PIS at Weeks 2, 4, 6, 10, 14, 18, 22, and 26
- Proportion of responders at Weeks 6, 10, 14, 18, 22, 26, and any of those weeks, where a responder is defined as having:
    - a) Clinically meaningful 4-week average morning ItchRO(Obs) severity decrease from baseline AND a normalization or reduction from baseline in sBA of ≥75%, or average sBA level <102 µmol/L
    - b) Clinically meaningful 4-week average morning ItchRO(Obs) severity decrease from baseline
    - c) a normalization or reduction from baseline in sBA of ≥75%, or average sBA level <102 µmol/L
- Number and proportion of participants achieving morning ItchRO(Obs) severity scores ≤1 for more than 50% of the time from baseline to Week 6, Weeks 7-10, 11-14, 15-18, 19-22, 23-26, and 15-26
- Mean change over time in quality of life as measured by the PedsQL
- Mean change from baseline in average sleep disturbance scores over time (4-week average score at Weeks 6, 10, 14, 18, 22, and 26), based on sleep-related questions on the EDQ(Obs) morning, PedsQL (parent), and PedsQL (child) questionnaires, individually

- Proportion of "no" answers for sleep-related questions on the ItchRO(Obs) morning, ItchRO(Obs) evening, and ItchRO(Pt) morning, individually, at the participant level for baseline to Week 6, Weeks 7-10, 11-14, 15-18, 19-22, 23-26, 15-26, and overall
- Proportion of participants with improvement from baseline (i.e., change from baseline <0) in EDQ(Obs) sleep score at Weeks 6, 10, 14, 18, 22, and 26 (using 4-week average scores)
- Proportion of EDQ sleep scores ≤2 at the participant level post-baseline through Week 26
- Mean change from baseline in height z-score and weight z-score at Weeks 2, 4, 6, 10, 14, 18, 22, and 26
- Proportion of participants who experience an improvement from baseline in height z-score or weight z-score (i.e., change from baseline >0) at Weeks 2, 4, 6, 10, 14, 18, 22, and 26
- Mean, change, and % change from baseline in total sBA level to Weeks 2, 6, 10, 14, 18, 22, and 26
- Mean change from baseline in bile acid subspecies, C4, FGF-19, and autotaxin to Weeks 10, 18, and 26
- Time to liver-associated events (PEBD surgery, listing for liver transplantation, liver decompensation [hepatic encephalopathy, variceal bleeding, ascites, and spontaneous bacterial peritonitis] events, HCC, death)
- Correlation analyses between sBA and pruritus severity

The ItchRO(Obs) frequency scores will be calculated using the same approach as the ItchRO(Obs) severity scores.

ItchRO(Pt), EDQ(Obs), and EDQ(Pt) will be scored and analyzed in the same manner as ItchRO(Obs) noted above.

Height and weight z-scores are based on a subject's gender and age at each scheduled visit. For subjects less than 24 months of age, the World Health Organization (WHO) growth charts are recommended by the Centers for Disease Control (CDC) and will be used to derive z-scores. For subjects at least 24 months of age, the CDC growth charts will be used to derive z-scores.

Severity scores for the EDQ are based on the EDQ(Obs) question of "How bad was your child's itch at its worst?" and on the EDQ(Pt) question of "How bad was your itch at its worst." "Non-severity" scores are based on all other EDQ questions.

The additional questions included in the ItchRO instrument that are not scored will be presented in subject listings.

To assess possible improvements in sleep, sleep-related questions from the ItchRO, EDQ, and PedsQL questionnaires will be identified and summarized. For ItchRO(Obs), quality of sleep information is collected on the morning and evening diary as it relates to contributing to the observer-rated severity score (1=Yes; 0=No). For ItchRO(Pt), the morning diary asks 2 sleep-related questions: "Did feeling itchy make it hard to fall asleep last night?" and "Did feeling itchy wake you up last night?" (1=Yes; 0=No). For EDQ(Obs), quality of sleep is assessed through the question, "Because of itch, my child had trouble staying asleep," which is scored on a 5-point scale. For the ItchRO and EDQ sleep-related questions, 4-week (morning and/or evening) average scores will be derived.

For PedsQL (parent and subject reports), quality of sleep is assessed by scoring difficulty in falling asleep, and/or difficulty/trouble in sleeping (mostly) through the night, on 3-point scales for young child subject reports or 5-point scales for all others.

The SAP will describe the subscales/domains, including the scoring algorithms, for each of the PedsQL questionnaires. Liver biochemistry (e.g., TSB, ALT), sBA normalization levels, along with definitions of elevated liver biochemistry levels and liver disease severity scores, will be covered in the SAP.

Analysis for Exploratory Efficacy Endpoints

An MMRM model with the proportion of days with a morning ItchRO(Obs) severity score ≤1 for each post-baseline time period as the outcome variable will be used to compare treatment differences for Weeks 0–6, 7–10, 11–14, 15–18, 19–22, and 23–26. The MMRM model will include the fixed, categorical effects of treatment group, time period, and treatment group-by-time period interaction as well as the continuous, fixed covariates of baseline 4-week average morning ItchRO(Obs) severity scores and the baseline score-by-time period interaction. Least squares means will be calculated for each treatment group for each post-baseline time period in the model. The difference between maralixibat and placebo change from baseline in the outcome variable will be estimated, with the corresponding 2-sided 95% CI constructed for each time period. Change from baseline LS means with standard error, 95% CI for the LS means, p-value for testing if the LS mean is zero, LS mean difference between treatment groups (maralixibat minus placebo) with standard error, 95% CI for the LS mean difference, and p-value for testing if the treatment LS means are equal will be calculated for each time period. In addition, an analysis of covariance (ANCOVA) model, with the proportion of days with morning ItchRO(Obs) severity score ≤1 for the last 12-weeks of the study (i.e., Weeks 15 to 26) as the outcome variable, treatment group as a factor, and baseline 4-week average morning ItchRO(Obs) severity scores as a covariate, will be used to compare the treatment difference for the proportion of days with morning ItchRO(Obs) severity score ≤1 for the last 12 weeks of the study. The proportion of days with morning ItchRO(Obs) severity score ≤1 for each subject will also be summarized by treatment group for each post-baseline time period (i.e., baseline to Week 6, Weeks 7–10, Weeks 11–14, Weeks 15–18, Weeks 19–22, Weeks 23–26, and Week 15–26.

Analysis similar to that described for the primary efficacy endpoint (i.e., MMRM) will be performed for each exploratory endpoint that is based on change from baseline values on

continuous variables. The repeated measures in the MMRM model will include the post-baseline visits as specified for each above-listed endpoint, with change from baseline value as the dependent variable. The MMRM model will include the fixed, categorical effects of treatment group, visit, and treatment group-by-visit interaction as well as the continuous, fixed covariates of baseline and the baseline-by-visit interaction. Least squares means will be calculated for each treatment group for each post-baseline visit in the model. The difference between maralixibat and placebo change from baseline in each outcome variable will be estimated, with the corresponding 2-sided 95% CI constructed for each visit. Change from baseline LS means with standard error, 95% CI for the LS means, p-value for testing if the LS mean is zero, LS mean difference between treatment groups (maralixibat minus placebo) with standard error, 95% CI for the LS mean difference, and p-value for testing if the treatment LS means are equal will be calculated for each visit. For the ItchRO and EDQ endpoints that compare treatment groups across the last 12 weeks of the study, the same methods as described for the primary analysis of the primary efficacy endpoint will be applied. The number of observations, mean, 95% CI on the mean, standard deviation, median, minimum, and maximum on both the observed and change from baseline values will also be summarized by treatment group for each visit including the end of study visit (Week 26/EOT).

For responder-type endpoints, the number and proportion of subjects that are considered a "responder" will be summarized by treatment group for each visit, including the end of study visit (Week 26/EOT). Barnard's exact test will be used to calculate the p-value for the difference between treatment groups at each study visit.

The number and proportion of subjects in each 5-point scale (0 to 4) of the CSS, CIS, and PIS, as well as the change from baseline in scale, will be summarized by treatment group for each visit including the end of study visit (Week 25/EOT). The Cochran-Mantel-Haenszel (CMH) test will be applied to test for treatment difference in the change from baseline scores at each study visit. Quality of sleep as assessed by the PedsQL parent and subject questionnaires will use similar analysis methods.

Kaplan-Meier curves will be used to estimate the distribution of time to liver-associated events (PEBD surgery, listing for liver transplantation, liver decompensation [hepatic encephalopathy, variceal bleeding, ascites, and spontaneous bacterial peritonitis] events, HCC, and death) for each treatment group. The median and other quartiles for time to liver-associated events, along with 2-sided 95% CIs, will be estimated based on the Kaplan-Meier method. Both a log-rank test and a Wilcoxon test will be used to test for treatment differences.

All tests will be performed as 2-sided tests at the 0.05 level of significance.

Sensitivity analyses will <u>not</u> be performed on the exploratory efficacy endpoints.

9.11 Safety and Tolerability Analyses

All safety analyses will be performed on the safety population, defined as all subjects who were randomized and received at least 1 dose of study medication. Analyses will be

conducted separately in participants with PFIC2 and participants with PFIC (PFIC1, PFIC2, PFIC3, PFIC4, PFIC5 and PFIC6).

For all safety and tolerability analyses, subjects will be analyzed by the treatment received and the overall active treatment group. For active treatment, the dose of IP received at the end of the dose-escalation period, or the last dose received if the subject discontinued from the IP during the dose escalation period, will be used.

Safety data collected at the baseline visit (Visit 1/Day 0) or the last preceding visit if not collected at Visit 1 will be used as the baseline value for safety analyses.

Safety measures including AEs, clinical laboratory values, physical examination findings (including body weight, height, and body mass index [BMI]), vital signs, ECGs, ultrasound liver imaging, and concomitant treatment usage will be summarized descriptively. For quantitative variables, descriptive statistics including number of observations, mean, median, standard deviation, minimum, and maximum will be presented for observed and change from baseline values at each study visit. Qualitative variables will be summarized using counts and percentages.

All safety and tolerability data will be presented in subject listings.

9.11.1 Adverse Events

In general, TEAEs are defined as AEs that start or deteriorate on or after the first dose of study medication and no later than 7 days following the last dose of study medication (for subjects not participating in the extension study) or reported through the Week 26/EOT visit (for subjects participating in the extension study). For any subjects who die during the study and the date of death is between the date of first dose of study medication and the date of study discontinuation (as entered by the site), inclusive, all AEs (including those resulting in death) that occur during the study will be considered as TEAEs irrespective of the last dose and will be included in the TEAE summaries. All summaries of AEs will be based on TEAEs unless specified otherwise.

Adverse events of special interest (AESIs) will be outlined in the SAP and summarized.

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The number and proportion of subjects who experience the event according to MedDRA system organ class (SOC) and preferred term will be presented by treatment group. Treatment-emergent AEs will be further summarized by severity and relationship to study medication. Adverse events related to study medication, AEs leading to withdrawal, SAEs, and deaths will be similarly summarized.

Subject listings of AEs will include the dose of study medication at the onset of the event.

9.11.2 Vital Signs and Weight/Height Measurements

Vital signs (temperature, systolic and diastolic blood pressure, heart rate, and respiratory rate), weight, height, and BMI will be summarized descriptively by study visit and treatment group as observed and change from baseline values. Weight and height measurements will also be summarized as a z-score for a subject's age and gender. Potentially clinically important (PCI) values for select vital signs results will be defined in the SAP. The number and proportion of subjects with PCI values will be presented by treatment group.

9.11.3 Laboratory Assessments

Safety laboratory test results will be summarized descriptively by study visit and treatment group as observed and change from baseline values. The number and proportion of subjects with clinical laboratory values below, within, or above the normal range by time point and in relation to baseline will be tabulated for each safety laboratory analyte by treatment group in a shift table. Potentially clinically important laboratory ranges will be defined in the SAP for select laboratory tests. The number and proportion of subjects with PCI laboratory values will be presented by treatment group.

Efficacy laboratory tests (i.e., total sBA, sBA subspecies, ALT, total and conjugated bilirubin, C4, FGF-19, autotaxin, FIB-4, and APRI) will not be included in safety summaries.

9.11.4 Electrocardiograms

Descriptive summaries will be presented for ECG measures of PR interval, QRS duration, and QT interval corrected using both Bazett's and Fridericia's formula (QTcB and QTcF). The number and proportion of subjects with normal and abnormal ECG results will also be summarized. These summaries will be presented by study visit and treatment group.

Potentially clinically important values for ECG results will be defined in the SAP. The number and proportion of subjects with potentially clinically important values will be presented by treatment group.

9.11.5 Liver Ultrasound

Liver ultrasound (or MRI) results will be presented in subject listings.

9.11.6 Concomitant Treatments

Prior and concomitant treatments will be summarized descriptively by treatment group using the number and proportion of subjects.

Prior medications will be presented separately from concomitant treatments. Medications that started prior to the first dose of study medication will be considered prior medications whether or not they were stopped prior to the first dose of study medication. Any medications continuing or starting after the first dose of study medication will be considered to be concomitant. If a medication starts prior to the first dose of study medication and

continues after the first dose of study medication, it will be considered both prior and concomitant.

Medications will be coded using the World Health Organization Drug Dictionary (WHO-DD).

9.12 Other Analyses

9.12.1 Healthcare Utilization

Healthcare resource utilization between baseline and Week 26 will be analyzed. Healthcare resource utilization variables include the number of hospitalizations, emergency room visits, the length of stay for hospitalization (days), surgeries and procedures related to the subject's PFIC condition, and the number of days the caregiver missed from work due to healthcare resource utilization events. Descriptive statistics including number of observations, mean, 95% CI on the mean, median, minimum, and maximum will be presented by visit for continuous variables. For categorical variables, the number and proportion of subjects will be presented.

9.12.2 Pharmacokinetic Analyses

Systemic concentrations of maralixibat in plasma will be determined at pre-dose and at approximately 2.5 hours after the morning dose at Week 10 and Week 26 (EOT/ET). Summary statistics (number of observations, mean, standard deviation, coefficient of variation, median, minimum, maximum, and geometric mean) will be determined for maralixibat concentrations at each nominal time point and sampling time.

10 SPONSOR'S AND INVESTIGATOR'S RESPONSIBILITIES

This study is conducted in accordance with current applicable regulations, ICH, EU Directive 2001/20/EC and its updates, and local ethical and legal requirements. Compliance with these regulations and guidelines also constitutes compliance with the ethical principles described in the Declaration of Helsinki.

The name and address of each third-party vendor (e.g., CRO) used in this study will be maintained in the investigator's and sponsor's files, as appropriate.

10.1 Sponsor's Responsibilities

10.1.1 Good Clinical Practice Compliance

The study sponsor and any third party to whom aspects of the study management or monitoring have been delegated will undertake their assigned roles for this study in compliance with all applicable industry regulations, ICH GCP Guideline E6 (1996) and E6 R2 (2017), EU Directive 2001/20/EC, as well as all applicable national and local laws and regulations.

Visits to sites are conducted by representatives of the study sponsor and/or the company organizing/managing the research on behalf of the sponsor to inspect study data, subjects' medical records and CRFs in accordance with current GCP and the respective local and (inter)national government regulations and guidelines. Records and data may additionally be reviewed by auditors or by regulatory authorities.

The sponsor ensures that local regulatory authority requirements are met before the start of the study. The sponsor (or a nominated designee) is responsible for the preparation, submission, and confirmation of receipt of any regulatory authority approvals required prior to release of study medication for shipment to the site.

10.1.2 Indemnity/Liability and Insurance

The sponsor ensures that suitable clinical study insurance coverage is in place prior to the start of the study.

If appropriate, a copy of the indemnity document is supplied to the investigator before study initiation, per local country guidelines.

10.1.3 Public Posting of Study Information

The sponsor is responsible for posting appropriate study information on applicable websites. Information included in clinical study registries may include participating investigators' names and contact information.

The timing for study registration and results summary posting must be in accordance with applicable local and national requirements.

10.1.4 Study Suspension, Termination, and Completion

The sponsor may suspend or terminate the study, or part of the study, at any time for any reason. If the study is suspended or terminated, the sponsor will ensure that applicable sites, regulatory agencies and IRBs/ECs are notified as appropriate. Additionally, the discontinuation of a registered clinical study which has been posted to a designated public website will be updated accordingly.

Should the study be temporarily suspended due to safety concerns, re-initiation will follow local laws and regulations (see Section 9.5).

The study may be terminated because of the following: the manifestation of an unjustifiable risk, a relevant toxicity with an unfavorable change to the benefit-risk ratio, a change in scientific knowledge (e.g., another toxicological or clinical study) with a negative impact on the evaluation of the benefit-risk ratio, a request or order from the competent authorities, the inability to complete recruitment within the prescribed timeframe, or the withdrawal of the authorization to manufacture or import the investigative product. In addition, interruption or premature termination of a clinical study can result if a study site is deemed unsuitable (by

the sponsor, authority, or Ethics Committee) or deviations from the approved SAP or the achievement or non-achievement of statistical objectives occur.

10.2 Investigator's Responsibilities

10.2.1 Good Clinical Practice Compliance

The investigator must undertake to perform the study in accordance with ICH GCP Guideline E6 (1996) and E6 R2 (2017), EU Directive 2001/20/EC, and applicable regulatory requirements and guidelines.

It is the investigator's responsibility to ensure that adequate time and appropriately trained resources are available at the site prior to commitment to participate in this study. The investigator should also be able to estimate or demonstrate a potential for recruiting the required number of suitable subjects within the agreed recruitment period.

The investigator will maintain a list of appropriately qualified persons to whom the investigator has delegated significant study-related tasks, and shall, upon request of the sponsor, provide documented evidence of any licenses and certifications necessary to demonstrate such qualification. Curriculum vitae for investigators and sub-investigators are provided to the study sponsor (or designee) before starting the study.

If a potential research subject has a primary care physician, the investigator should, with the subject's consent, inform them of the subject's participation in the study.

10.2.2 Protocol Adherence and Investigator Agreement

The investigator and any sub-investigators must adhere to the protocol as detailed in this document. The investigator is responsible for enrolling only those subjects who have met protocol eligibility criteria. Investigators are required to sign an investigator agreement to confirm acceptance and willingness to comply with the study protocol.

If the investigator suspends or terminates the study at their site, the investigator will promptly inform the sponsor and the IRB/EC and provide them with a detailed written explanation. Upon study completion, the investigator will provide the sponsor, IRB/EC, and regulatory agency with final reports and summaries as required by (inter)national regulations.

10.2.3 Documentation and Retention of Records

10.2.3.1 Case Report Forms

Electronic CRFs should be handled in accordance with instructions from the sponsor.

The investigator is responsible for maintaining adequate and accurate medical records from which accurate information is recorded onto CRFs, which have been designed to record all observations and other data pertinent to the clinical investigation. CRFs must be completed by the investigator or designee as stated in the site delegation log.

All data submitted in the electronic CRF must be reviewed, approved and signed for by the investigator. All data, except data entered directly into the eDiary, will have separate source documentation.

The clinical research associate (CRA)/study monitor will verify the contents against the source data per the monitoring plan. If the data are unclear or contradictory, queries are sent for corrections or verification of data.

10.2.3.2 Recording, Access, and Retention of Source Data and Study Documents

Original source data to be reviewed during this study will include but are not limited to subject's medical file, daily dosing records and clinical laboratory reports.

The investigator must permit authorized representatives of the sponsor; the respective national, local, or foreign regulatory authorities; the IRB/EC; and auditors to inspect facilities and to have direct access to original source records relevant to this study, regardless of media.

The consent form includes a statement by which the subject agrees to the monitor/auditor from the sponsor or its representatives, national or local regulatory authorities, or the IRB/EC, having access to source data. Non-study site personnel will not disclose any personal information or personal medical information.

Records must be made available within reasonable times for inspection and duplication, if required, by a properly authorized representative of any regulatory agency (e.g., the US Food and Drug Administration, European Medicines Agency, UK Medicines and Healthcare products Regulatory Agency) or an auditor.

Essential documents must be maintained according to ICH GCP requirements and in accordance with local country-specific regulations and may not be destroyed without written permission from the sponsor.

10.2.3.3 Audit/Inspection

To ensure compliance with relevant regulations, data generated by this study must be available for inspection upon request by representatives of, for example, the US FDA (as well as other US national and local regulatory authorities), the European Medicines Agency, the Medicines and Healthcare products Regulatory Agency, other regulatory authorities, the sponsor or its representatives, and the IRB/EC for each site.

10.2.3.4 Financial Disclosure

The investigator is required to disclose any financial arrangement during the study and for 1 year after, whereby the outcome of the study could be influenced by the value of the compensation for conducting the study, or other payments the investigator received from the sponsor. The following information is collected: any significant payments from the sponsor or subsidiaries such as a grant to fund ongoing research, compensation in the form of

equipment, retainer for ongoing consultation or honoraria; any proprietary interest in IP; any significant equity interest in the sponsor or subsidiaries as defined in 21 CFR 54.2(b).

10.3 Ethical Considerations

10.3.1 Informed Consent/Assent

It is the responsibility of the investigator or designee to obtain written informed consent (or assent as applicable) from all study subjects prior to any study-related procedures including screening assessments. All consent and assent documentation must be in accordance with applicable regulations and GCP. Each subject or the subject's legally authorized representative, as applicable, is requested to sign and date the subject informed consent form or a certified translation if applicable, after the subject has received and read (or been read) the written subject information and received an explanation of what the study involves, including but not limited to: the objectives, potential benefits and risk, inconveniences, and the subject's rights and responsibilities. A copy of the informed consent and assent documentation (i.e., a complete set of subject information sheets and fully executed signature pages) must be given to the subject or the subject's legally authorized representative, as applicable. Signed consent and assent forms must remain in each subject's study file and must be available for verification at any time.

The PI provides the sponsor with a copy of the consent form (or assent as applicable) that was reviewed by the IRB/EC and received their favorable opinion/approval. A copy of the IRB/EC's written favorable opinion/approval of these documents must be provided to the sponsor prior to the start of the study unless it is agreed to and documented (abiding by regulatory guidelines and national provisions) prior to study start that another party (i.e., sponsor or coordinating PI) is responsible for this action.

10.3.2 Institutional Review Board or Ethics Committee

This protocol, the informed consent document (approved by the sponsor or their designee), relevant supporting information and all types of subject recruitment information will be submitted to the IRB/EC for review, and all must be approved prior to site initiation. Study medication supplies will not be released until written IRB/EC approval has been received.

Prior to implementing changes in the study, the sponsor and the IRB/EC must approve any amendments to the protocol and revisions of all informed consent documents, unless there is a subject safety issue.

The IRB/EC will be kept apprised of the progress of the study and of any changes made to the protocol. The IRB/EC will also be informed of any serious and significant AEs.

10.4 Privacy and Confidentiality

All US-based sites and laboratories or entities providing support for this study, must, where applicable, comply with the Health Insurance Portability and Accountability Act (HIPAA) of

1996. A site that is not a covered entity as defined by HIPAA must provide documentation of this fact.

The confidentiality of subject records in EU countries will be protected in accordance to the General Data Protection Regulation guideline.

The confidentiality of records that may be able to identify subjects will be protected in accordance with applicable laws, regulations, and guidelines.

After subjects have consented to take part in the study, the sponsor and/or its representatives may review medical records and data collected during the study. These records and data may, in addition, be reviewed by others including the following: independent auditors who validate the data on behalf of the sponsor; third parties with whom the sponsor may develop, register, or market maralixibat; national or local regulatory authorities; and the IRBs/ECs which gave approval for the study to proceed. The sponsor and/or its representatives accessing the records and data will take all reasonable precautions in accordance with applicable laws, regulations, and guidelines to maintain the confidentiality of subjects' identities.

Subjects are assigned a unique identifying number; however, their initials and date of birth may also be collected, if permitted under local laws governing privacy.

The results of studies – containing subjects' unique identifying number, relevant medical records, and possibly initials and dates of birth, where allowed per local law, may be transferred to, and used in, other countries which may not afford the same level of protection that applies within the countries where this study is conducted. The purpose of any such transfer would include: to support regulatory submissions, to conduct new data analyses to publish or present the study results, or to answer questions asked by regulatory or health authorities.

10.5 Study Results/Publication Policy

The sponsor will endeavor to publish the results of all qualifying, applicable, and covered studies according to external guidelines in a timely manner regardless of whether the outcomes are perceived as positive, neutral, or negative.

All publications relating to sponsor products or projects must undergo appropriate technical and intellectual property review, with sponsor agreement to publish prior to release of information.

The term "publication" refers to any public disclosure including original research articles, review articles, oral presentations, abstracts and posters at medical congresses, journal supplements, letters to the editor, invited lectures, opinion pieces, book chapters, electronic postings on medical/scientific websites, or other disclosure of the study results, in printed, electronic, oral or other form.

11 REFERENCES

Al Mehaidib A, Al Shahrani A. 1381 Progressive familial intrahepatic cholestasis in Arabs. J Hepatol 2013;58:S555-6.

Alissa FT, Jaffe R, Shneider BL. Update on progressive familial intrahepatic cholestasis. J Pediatr Gastroenterol Nutr 2008;46:241-52.

Arnell H, Papadogiannakis N, Zemack H, et al. Follow-up in children with progressive familial intrahepatic cholestasis after partial external biliary diversion. J Pediatr Gastroenterol Nutr 2010;51:494-9.

Cnaan A, Laird NM, Slasor P. Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. Stat Med 1997;16, 2349-80.

Davit-Spraul A, Gonzales E, Baussan C, et al. Progressive familial intrahepatic cholestasis. Orphanet J Rare Dis 2009;4:1.

Emerick KM, Elias MS, Melin-Aldana H, et al. Bile composition in Alagille syndrome and PFIC patients having partial external biliary diversion. BMC Gastroenterol 2008;8:47.

Ethical Considerations for Clinical Trials on Medicinal Products Conducted with Minors. 2017. Accessed at: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/2017_09_18_ethical_consid_ct_with_minors.pdf

European Medicines Agency. 2010. Guideline on missing data in confirmatory clinical trials.

Halaweish I, Chwals WJ. Long-term outcome after partial external biliary diversion for progressive familial intrahepatic cholestasis. J Pediatr Surg 2010;45:934-7.

Hounnou G, Destrieux C, Desme J, et al. Anatomical study of the length of the human intestine. Surg Radiol Anat 2002;24:290-4.

Jericho H. Bile acid pool dynamics in progressive familial intrahepatic cholestasis with partial external bile diversion. J Pediatr Gastroenterol Nutr 2015;60:368-74.

Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 1997;53:983-97.

Little R, Rubin D. Statistical analysis with missing data. New York: John Wiley: 1987.

Little R, Yau L. Intent-to-treat analysis for longitudinal studies with drop-outs, Biometrics 1996;52:1324-33.

Mallinckrodt CH, Clark WS, David SR. Accounting for dropout bias using mixed-effects models. J Biopharm Stat 2001;11:9-21.

Mallinckrodt CH, Lane PW, Schnell D, et al. Recommendations for the primary analysis of continuous endpoints in longitudinal clinical trials. Drug Information Journal 2008;42:303-19.

Mallinckrodt CH, Roger J, Chuang-Stein C, et al. Recent developments in the prevention and treatment of missing data. Drug Information Journal 2013;48:68-80.

Molenberghs G, Kenward MG. Missing data in clinical studies. New York: Wiley:2007.

Molenberghs G, Thijs H, Jansen I, et al. Analyzing incomplete longitudinal clinical trial data. Biostatistics 2004;5:445-64.

Pawlikowska L, Strautnieks S, Jankowska I, et al. Differences in presentation and progression between severe FIC1 and BSEP deficiencies. J Hepatol 2010;53:170-8.

Ratitch B, O'Kelly M. 2011. Implementation of pattern-mixture models using standard SAS/STAT procedures. Proceedings of PharmaSUG 2011 (Pharmaceutical Industry SAS Users Group), SP04, Nashville.

Sambrotta M. Mutations in TJP2 cause progressive cholestatic liver disease. Nat Genet 2014;46:326-8.

Schafer JL. Analysis of incomplete multivariate data. New Yori: Chapman & Hall:1997.

Schukfeh N, Metzelder ML, Petersen C, et al. Normalization of serum bile acids after partial external biliary diversion indicates an excellent long-term outcome in children with progressive familial intrahepatic cholestasis. J Pediatr Surg 2012;47:501-5.

Shneider BL, Magee JC, Bezerra JA, et al. Efficacy of fat-soluble vitamin supplementation in infants with biliary atresia. Pediatrics 2012;130:607-14.

Siddiqui O, Hung HM, O'Neill R. MMRM vs. LOCF: a comprehensive comparison based on simulation study and 25 NDA datasets. J Biopharm Stat 2009;19:227-46.

Stapelbroek JM, van Erpecum KJ, Klomp LWJ, et al. Liver disease associated with canalicular transport defects: current and future therapies. J Hepatol 2010;52:258-71.

Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. Medical Care 2001;39:800-12.

Verbeke G, Molenberghs G. Linear mixed models for longitudinal data. New York: Springer: 2000.

Weaver LT, Austin S, Cole TJ. Small intestinal length: a factor essential for gut adaptation. Gut 1991;32:1321-3.

Whitington PF, Whitington GL. Partial external diversion of bile for the treatment of intractable pruritis associated with intrahepatic cholestasis. Gastroenterology 1988;95:130-6.

Yang H, Porte RJ, Verkade HJ, et al. Partial external biliary diversion in children with progressive familial intrahepatic cholestasis and Alagille disease. J Pediatr Gastroenterol Nutr 2009;49:216-21.

12 APPENDICES

Appendix 1 Clinical Outcomes Assessments

The following COAs will be utilized in this study:

Clinical Outcomes Assessment	Completed by*
ItchRO(Obs)™	Caregiver
ItchRO(Pt)™	Subjects aged ≥9 years
PedsQL Child Report	Subjects aged 8–12 years
PedsQL Teenager Report	Subjects aged 13–18 years
Parent PedsQL Report	Caregivers using age-appropriate report (Report for Infants, Toddlers, Young Children, Children, or Teenagers)
PIS	Subjects aged ≥9 years
CIS	Caregiver
EDQ(Obs)	Caregivers for subjects aged <9 years
EDQ(Pt)	Subjects aged ≥9 years
PIC	Subjects aged ≥9 years
CIC	Caregiver

CIS=Caregiver Impression of Severity of Pruritus; ItchRO(Obs)™=Itch Reported Outcome observer instrument; ItchRO(Pt)™=Itch Reported Outcome patient instrument; PedsQL=Pediatric Quality of Life Inventory; PIS=Patient Impression of Severity; EDQ: Exploratory Diary Questionnaire; PIC: Patient impression of Change; CIC: Caregiver Impression of Change

* For the ItchRO and EDQ, the subject's age at the screening visit (Visit 0) will be used for the determination of the appropriate module to be used for the study.
For the PedQL and PIS, the subject's age at the screening visit (Visit 0) will be used for the determination of the appropriate module to be used for the study.
The same module will be used for the duration of the study regardless of subsequent birthdays throughout the study.
For the PIC, subjects aged ≥9 years at Visit 9/EOT should complete the diary.

Appendix 2 Observer Itch Reported Outcome Instrument ItchRO(Obs)

ItchRO(Obs): Morning Diary

Based on observations or what your child told you about his/her itching, how severe were your child's itch-related symptoms (rubbing, scratching, skin damage, sleep disturbances or irritability) from when he/she went to bed last night until he/she woke up this morning?

Select one response below.

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very severe
- ☐ I don't know

Below, please select all that contributed to your answer.

- ☐ Child reported itching
- ☐ Observed difficulty falling asleep or staying asleep (sleep disturbance)
- ☐ Observed rubbing or scratching
- ☐ Observed new or worsening marks on the skin due to rubbing or scratching
- ☐ Observed fussiness or irritability

While you were observing your child from when he/she went to bed last night until he/she woke up this morning, how much of the time was your child rubbing or scratching?

Select one response below.

- ☐ None
- ☐ A little bit of the time
- ☐ Some of the time
- ☐ Most of the time
- ☐ Almost all of the time/constantly
- ☐ I don't know

ItchRO(Obs): Evening Diary

Based on observations or what your child told you about his/her itching, how severe were your child's itch-related symptoms (rubbing, scratching, skin damage, sleep disturbances or irritability) from the time he/she woke up this morning until he/she went to bed?

Select one response below.

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very severe
- ☐ I don't know

Below, please select all that contributed to your answer.

- ☐ Child reported itching
- ☐ Observed difficulty falling asleep or staying asleep (sleep disturbance)
- ☐ Observed rubbing or scratching
- ☐ Observed new or worsening marks on the skin due to rubbing or scratching
- ☐ Observed fussiness or irritability

While you were observing your child from the time he/she woke up this morning until he/she went to bed, how much of the time was your child rubbing or scratching?

Select one response below.

- ☐ None
- ☐ A little bit of the time
- ☐ Some of the time
- ☐ Most of the time
- ☐ Almost all of the time/constantly
- ☐ I don't know

Appendix 3 Patient Itch Reported Outcome Instrument ItchRO(Pt)

ItchRO(Pt): Morning Diary

Think about whether itching kept you awake or woke you up last night. Think about whether you felt like rubbing or scratching.

How itchy did you feel last night after you went to bed until you woke up this morning?

Select one response below.

- ☐ I didn't feel itchy
- ☐ I felt a little bit itchy
- ☐ I felt pretty itchy
- ☐ I felt very itchy
- ☐ I felt very, very itchy

Did feeling itchy make it hard to fall asleep last night?

- ☐ Yes
- ☐ No

Did feeling itchy wake you up last night?

- ☐ Yes
- ☐ No

How much of the night did feeling itchy make you rub or scratch?

- ☐ None of the night
- ☐ Just a tiny bit of the night
- ☐ Half of the night
- ☐ Most of the night
- ☐ All night

ItchRO(Pt): Evening Diary

Think about how itchy you were all day. Think about whether you felt like rubbing or scratching.

How itchy were you all day today from the time when you woke up until now?

Select one response below.

- ☐ I didn't feel itchy
- ☐ I felt a little bit itchy
- ☐ I felt pretty itchy
- ☐ I felt very itchy
- ☐ I felt very, very itchy

Did feeling itchy make you rub or scratch today?

- ☐ No
- ☐ Yes, but it left no marks
- ☐ Yes, and it left marks, but my skin wasn't red
- ☐ Yes, and it left red marks
- ☐ Yes, and my skin bled

How much of today did feeling itchy make you rub or scratch?

- ☐ None of the day
- ☐ Just a tiny bit of the day
- ☐ Half of the day
- ☐ Most of the day
- ☐ All day

Appendix 4 Pediatric Quality of Life Inventory (PedsQL)

Child's age at screening (Visit 0)	Questionnaires required to be completed at Day 0, Week 4 or 6[a], Week 10, Week 14, Week 18, Week 22, and Week 26 (EOT/ET)	Who completes?
1–12 months		
	Parent Report for Infants (1–12 months)	Parent
	Family Impact Module	Parent
13–24 months		
	Parent Report for Infants (13–24 months)	Parent
	Family Impact Module	Parent
2–4 years old		
	Parent Report for Toddlers (2–4)	Parent
	Multidimensional Fatigue Scale, Parent Report (2–4)	Parent
	Family Impact Module	Parent
5–7 years old		
	Parent Report for Young Children (5–7)	Parent
	Multidimensional Fatigue Scale, Parent Report for Young Children (5–7)	Parent
	Family Impact Module	Parent
8–12 years old		
	Parent Report for Children (8–12)	Parent
	QOL inventory for Children (8–12)	Subject
	Multidimensional Fatigue Scale, Parent Report for Children (8–12)	Parent
	Multidimensional Fatigue Scale, Child Report for ages (8–12)	Subject
	Family Impact Module	Parent
13–18 years old		
	Parent Report for Teenagers (13–18)	Parent
	QOL inventory for Teenagers (13–18)	Subject
	Multidimensional Fatigue Scale, Parent Report for Teenagers (13-18)	Parent
	Multidimensional Fatigue Scale, Teen Report for Children (13–18)	Subject
	Family Impact Module	Parent

Appendix 5 Impression of Severity Questionnaires

Patient Impression of Severity (PIS)

Q: Overall, how severe was your itching (pruritus) over the past 7 days?

Select one response below.

- ☐ I didn't feel itchy at all
- ☐ I felt a little bit itchy
- ☐ I felt pretty itchy
- ☐ I felt very itchy
- ☐ I felt very, very itchy

Caregiver Impression of Severity (CIS)

Q: Based on observations or what your child told you about his/her itching (pruritus), how severe were your child's itch-related symptoms over the past 7 days?

Select one response below.

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very severe

Appendix 6 Exploratory Diary Questionnaire

Observer Morning Diary

Instructions: Based on what you saw or what your child told you, think about your child's itching from when he/she went to bed last night until he/she woke up this morning.
1) It was hard for my child to stop scratching or rubbing. 1 – Never 2 – Rarely 3 – Sometimes 4 – Often 5 – Almost Always
2) My child scratched until he/she bled. 1 – Never 2 – Rarely 3 – Sometimes 4 – Often 5 – Almost Always
3) Because of itch, my child had trouble staying asleep. 1 – Never 2 – Rarely 3 – Sometimes 4 – Often 5 – Almost Always
4) How bad was your child's itch at its worst? 1 – No itch 2 – Mild 3 – Moderate 4 – Severe 5 – Very Severe

Observer Evening Diary

Instructions: Based on what you saw or what your child told you, think about your child's itching from when he/she woke up this morning until he/she went to bed.
1) It was hard for my child to stop scratching or rubbing. 1 – Never 2 – Rarely 3 – Sometimes 4 – Often 5 – Almost Always
2) My child scratched until he/she bled. 1 – Never 2 – Rarely 3 – Sometimes 4 – Often 5 – Almost Always
3) How bad was your child's itch at its worst? 1 – No itch 2 – Mild 3 – Moderate 4 – Severe 5 – Very Severe

Patient Morning Diary

Instructions: Think about your itching from when you went to bed last night until you woke up this morning.
1) How bad was your itch at its worst? 1 – No itch 2 – Mild 3 – Moderate 4 – Severe 5 – Very Severe

Patient Evening Diary

Instructions: Think about your itching from when you woke up this morning until you went to bed.
1) How often did you itch? 1 – Never 2 – Rarely 3 – Sometimes 4 – Often 5 – Almost Always
2) How bad was your itch at its worst? 1 – No itch 2 – Mild 3 – Moderate 4 – Severe 5 – Very Severe

Appendix 7 Clinician Scratch Score

This scoring scale was originally developed to assess pruritus before and after surgical intervention in children with ALGS and PFIC (Whitington and Whitington 1988).

The clinician will rate the subject's pruritus, as evidenced by scratching, according to the following scale:

Score	Description
0	None
1	Rubbing or mild scratching when undistracted
2	Active scratching without evident skin abrasions
3	Abrasion evident
4	Cutaneous mutilation, haemorrhage and scarring evident

Appendix 8 Patient Impression of Change (PIC)

The PIC is designed to assess the subject's perception of his/her itching at Week 26 (EOT) compared to his/her itching prior to the start of treatment with study drug. The PIC will be completed, by subjects who are 9 years of age or older at the Week 26 (EOT) visit (see Table 1)

The questionnaire is designed for self-administration and uses a 7-point scale in which 1 designates the best outcome and 7 designate the worst outcome.

PIC

How much has your itching changed, if at all, since you started this study?

☐ Much better (1)

☐ Better (2)

☐ A little better (3)

☐ No change (4)

☐ A little worse (5)

☐ Worse (6)

☐ Much worse (7)

Appendix 9 Caregiver Impression of Change (CIC)

The CIC is designed to assess the caregiver's perception of the subject's itch related symptoms and xanthoma severity at Week 26 (EOT) compared to his/her itch related symptoms and xanthoma severity prior to the start of treatment with study drug. The CIC will be completed by all caregivers at the Week 26 (EOT) visit (see Table 1).

The questionnaire is designed for self-administration and uses a 7-point scale in which 1 designates the best outcome and 7 designates the worst outcome.

CIC

How would you rate the change in your child's itch related symptoms since the start of the study?

- ☐ Much better (1)
- ☐ Better (2)
- ☐ A little better (3)
- ☐ No change (4)
- ☐ A little worse (5)
- ☐ Worse (6)
- ☐ Much worse (7)

How would you rate the change in your child's xanthoma severity since the start of the study?

☐ Much better (1)

☐ Better (2)

☐ A little better (3)

☐ No change (4)

☐ A little worse (5)

☐ Worse (6)

☐ Much worse (7)

Appendix 10 Management of Clinical Study Procedures and Participants during COVID-19 Pandemic or Other Force Majeure

This appendix provides guidance for participant safety and ongoing access to medical care and investigational product during the global novel coronavirus 2019-nCoV (COVID-19) pandemic. During this global public health crisis, pragmatic and harmonized actions are required to ensure the necessary flexibility and procedural simplifications that are needed to maintain the integrity of the clinical studies and to ensure the rights, safety, and wellbeing of study participants and the safety of clinical study staff.

The measures detailed here are implemented across Mirum studies on a temporary basis until the pandemic is considered resolved by governmental and public health organizations, as applicable.

Regardless of the guidance provided below, consideration should be given to public health advice at the study locations and individual benefit/risk in treatment decisions for participants at the study site during the pandemic (EMA 2020; FDA 2020). Logistical requirements should also be considered, such as the ability of participants to travel to the study site, accessibility of public transportation, site restrictions on clinical studies, etc.

The safety of the study participants is of primary importance, and risks of involvement in the study, in particular with added challenges due to COVID-19, will be weighed against anticipated benefit for the study participants and society.

Study Participation

Study sites may continue to recruit new participants, if deemed appropriate on benefit-risk assessment as described, and provided that ALL of the following activities to preserve study integrity can be met:

- Upon discussion with the site monitor, the study site has confirmed the ability to enroll and manage studies and participants effectively and in compliance with the protocol and this appendix.
- Upon discussion with the sponsor, medical monitor, or designee, the study site has confirmed that appropriate safety monitoring can take place, in compliance with the protocol and this appendix.
- Data will continue to be entered into the electronic Case Report Form (eCRF) and queries resolved in a timely manner.
- The site monitor is able to access the study site to perform onsite monitoring or is able to perform remote monitoring of data.

Conduct of Telephone/Virtual or Alternative Visits

Due to the current pandemic, it is conceivable that not all participant visit commitments can be fulfilled. If a participant or caregiver is unable or unwilling to attend a study visit or a study site is unable to permit onsite visits by participants and caregivers, adaptation of the onsite visit to a telephone visit, virtual visit, or use of an alternative location for assessment (including but not limited to local laboratories, family physician, home visit) is recommended to ensure continuity of participant care during the study (as an interim measure). Priority should be given to maintaining ongoing safety follow-up. Study sites should speak with their site monitor before performing a telephone or virtual visit so he or she may provide guidance regarding logistics that may need consideration. Study sites should contact the medical monitor if the participant must miss more than one onsite visit in succession because multiple incomplete visits may have the potential to impact evaluation of study endpoints as well the participant's safety.

The decision to accept protocol deviations for missed assessments is based on the risk to the participant of discontinuing study participation versus missed assessments, such as safety laboratory assessments. Where possible, a homecare visit should be organized to avoid missing visits or assessments.

Alternative Blood Sample Collection:

Where possible and permitted, homecare can be organized through qualified health care professionals, trained on the study protocol, to collect samples at the participant's home (or other agreed upon location). It is preferred that this alternative blood sample collection is performed in a manner to support delivery of samples to the central laboratory, but the analysis of samples may be performed at local laboratories, if necessary. If the analyses are performed at the local laboratories, the site will follow up with the participant to obtain the results and local laboratory reference ranges for data entry into EDC as soon as they are available.

Safety Reporting and Maralixibat Treatment Recommendations for COVID-19

The sponsor recognizes that COVID-19 presents an increased risk for all participants. Due to the potential impact of COVID-19 on multiple organ systems, the sponsor recommends the following dose modification and management plan for participants with confirmed or suspected COVID-19 while receiving treatment with maralixibat.

The following safety reporting guidelines are required:

All confirmed or suspected COVID-19–related adverse events (AEs) must be recorded in the eCRF and the Serious Adverse Event (SAE) Report Form if the event meets the seriousness criteria. SAEs are reportable to the sponsor immediately without undue delay after becoming aware of the event and no later than 24 hours of awareness. All study-drug interruptions or modifications must be recorded on the AE and drug administration eCRFs.

If an event is suspected to be COVID-19 infection, continuation of study drug, and any dose modifications should be assessed on a case-by-case basis, depending on the participant's health status, and discussed with the medical monitor if needed. The COVID-19 infection should be managed as per local treatment guidance and investigator's clinical judgment.

Drug-Drug Interaction Guidance Regarding Medication Used to Treat COVID-19

- The study protocol and Investigator's Brochure should be referenced for drug-drug interaction information. For further information the medical monitor can be contacted.

Guidance Regarding Vaccination against COVID-19

- Participants should be requested to contact the site upon planned vaccination against COVID-19. The vaccine should be entered in the eCRF as a concomitant medication and any new or worsened symptom or sign recorded as an adverse event and assessed for seriousness. Maralixibat is taken orally, is minimally absorbed with systemic levels below the lower limit of quantification, and binds with and inhibits the ileal bile acid transporter (IBAT) in the distal ileum. The vaccine is administered and acts systemically; hence, no drug-drug interaction is expected, and no need for unblinding is foreseen.

Documenting Alternative Contacts with Participants:

If an onsite visit is replaced by a telephone or virtual study visit, the following are guidelines for data capture:

- If the visit is listed as onsite but is performed remotely, data should be captured as per a normal visit (i.e., document the remote visit in the source documents and complete the relevant eCRF pages to capture the visit date, and any possible assessment that can be obtained, such as AEs, study drug administration, and/or concomitant medications and any additional safety or efficacy information). All assessments that cannot be performed should be marked as not done in the CRF. Any protocol deviations related to COVID-19 should be recorded in the clinic notes, prefixed COVID-19.

- Visits that require procedures that cannot be performed remotely should be discussed with the site monitor because this may impact efficacy or safety analyses and should be documented as a protocol deviation due to COVID-19.

Site-to-Participant Drug Shipment Instructions during Pandemic Containment

If a participant is unable to go to the study site because of pandemic containment, the study site may ship the study drug to the home of the participant (or other agreed location) following approval by the sponsor.

For such shipment, the following conditions must be met:

- The sponsor is responsible for delivery of the study drug to the study site. Organization of shipments from the site to the participant will be the responsibility of the study site.

- The participant or caregiver is informed about the shipment method, confirms the address for receipt of the drug, and agrees that his or her personal information (i.e., contact information) may be given to a professional carrier.
- The investigator or designee (e.g., pharmacy) securely packages the drug for shipment.
- A professional carrier is used by the investigator or designee to ship the drug securely and maintain chain of custody, with evidence provided. Maralixibat should be stored and shipped as directed in the Pharmacy Manual. Shipments that are expected to be completed within 48 hours from pickup to delivery do not generally require any specific temperature monitoring.
- To respect participant confidentiality, the carrier should be given only the contact information of the recipient. The sponsor should not receive any personal information about the participant.
- A procedure is defined with the carrier to immediately upon delivery confirm the receipt of the drug by the recipient and that it is received in good condition.
- The site contacts the recipient to confirm the receipt of the drug and to confirm it is delivered in the expected conditions (e.g., not leaking, not broken) and gives instructions about the drug administration and storage. If the drug is delivered and is not within the expected conditions, the site will give instructions to the participant about the next steps.
- The investigator or designee completes the drug accountability forms with each shipment made directly to a participant.
- All documentation (originals/copies, as applicable) related to the site-to-participant drug shipment should be filed in the investigator site file, including the list of the medication being delivered, the quantity involved, and documentation of receipt by the study participant.
- Participants should retain unused study drug and containers and return them to the investigator the next time they visit the investigator site

COVID-19 Specific Remote Source Data Verification

The source documents/source data considered for remote source data verification (rSDV) are those related to the primary endpoint, AEs/SAEs, important medical events, or the reasons for exclusion of a participant from the study.

Remote access to source documents/source data for monitoring purposes may only take place in justified exceptional cases and only to the extent strictly necessary (i.e., only when direct access to study source data is not available due to COVID-19 pandemic restrictions).

The following 3 options for source data reconciliation without the study monitor being physically present at the study site may be used, based on the site-specific source documentation/source data:

- The study site will provide the study monitor, under the responsibility of the investigator, with copies of the source documents/source data in which personal

identifying information of the study participants and information pertaining to their privacy has been obscured or redacted.

- Under the responsibility of the investigator, the study site will grant the study monitor direct, controlled, remote access to the systems with which the source documents/source data are managed.

- The study site will grant the study monitor, under the responsibility of the investigator, passive access to the source documents/source data via live image transmission (e.g., sharing of the screen or image/sound transmission).

Depending on the option selected for rSDV, the full process to be followed and requirements needed to be met are described in the specific rSDV section of the clinical monitoring plan as well as in a corresponding guideline for the investigators.

In the case of live image transmission, this will take place exclusively via secured systems/ environments and systems/servers within European Economic Area/European Union.

In all cases, rSDV will be conducted exclusively by the authorized persons in accordance with the written informed consent of the study participants and the written agreement of the Principal Investigator (PI)/PI's institution.

Remote source data verification will take place in a protected environment (i.e., providing protection from unauthorized access in any form, including the use of privacy screens to prevent unauthorized viewing of source documents/source data), the source documents/source data reviewed by the study monitor will not be permanently stored by the study monitor, and if necessary, temporarily saved files (including ones automatically generated by the system) are permanently deleted at short notice.

The study monitor will securely destroy any copy of obscured/redacted documents, whether paper or electronic, as soon as they have been used for source data verification.

Appropriate corrective measures will be implemented in the event of technical difficulties or if the security of the transmission is no longer guaranteed.

The information and communication technology used by the sponsor and the study site for rSDV is designed in such a way that secure and General Data Protection Regulation–compliant transmission is guaranteed.

REFERENCES

United States Food and Drug Administration. FDA guidance on conduct of clinical trials of medical products during COVID-19 public health emergency. March 2020 (updated 30 August 2021). Available online at https://www.fda.gov/regulatory-information/search-fda-guidance-documents/fda-guidance-condu,ct-clinical-trials-medical-products-during-covid-19-public-health-emergency.

European Medicines Agency. Guidance on the management of clinical trials during the COVID-19 (coronavirus) pandemic. Version 3, 28 April 2020 (Updated Version 4, 4 February 2021). Available online at https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/guidanceclinicaltrials_covid19_en.pdf.