STATISTICAL ANALYSIS PLAN

Title:	A Phase 3 Randomized, Double-Blind, Placebo-Controlled Study of Adeno-Associated Virus Serotype 8-Mediated Gene Transfer of Glucose-6-Phosphatase in Patients with Glycogen Storage Disease Type Ia
Protocol:	DTX401-CL301
Investigational Product:	DTX401
Phase:	3
Sponsor:	Ultragenyx Pharmaceutical Inc. 60 Leveroni Court Novato, CA, USA 94949
Author:	PPD
Date:	5 April 2024
Version Number:	Version 1.0

TABLE OF CONTENTS

LIST OF ABBREVIATIONS 8
DOCUMENT REVISION HISTORY 10
1. INTRODUCTION AND OBJECTIVES OF ANALYSIS 11
1.1. Introduction 11
1.2. Objectives of Analysis 11
2. STUDY OBJECTIVE(S) 12
2.1. Estimands 15
3. STUDY DESIGN 19
3.1. Study Population 21
3.2. Dosage and Administration 22
3.3. Blinding and Randomization Methods 23
3.3.1. Blinding Strategy 23
3.3.2. Randomization 23
3.4. Potential Covariate(s) 23
3.5. Subgroups 23
3.6. Sample Size Considerations 23
3.7. Statistical Hypothesis 24
3.8. Multiple Comparisons/Multiplicity 24
3.9. Interim Analysis 25
3.10. Study Oversight Committee 25
3.10.1. Data Monitoring Committee 25
3.10.2. Central Independent Committee 26
4. DEFINITIONS 27
4.1. Analysis Study Periods 27
4.2. Analysis Treatment Groups 27
4.3. Baseline 27
4.4. Duration of exposure 27
4.5. Study Day 28
5. ANALYSIS SETS 29
5.1. Screened Set 29
5.2. Enrolled Set 29

5.3. Full Analysis Set (FAS) 29
5.4. Safety Analysis Set 29
5.5. Intention to Treat Analysis Set (ITT) 29
5.6. Modified Intention to Treat Analysis Set (mITT) 29
5.7. Per Protocol Set (PP) 29
6. STATISTICAL ANALYSIS 31
6.1. General Principles 31
6.2. Subject Accountability 31
6.3. Protocol Deviations and Intercurrent Events 32
6.4. Investigational Product Administration 32
6.4.1. Extent of Exposure 32
6.4.2. Measurement of Treatment Compliance 33
6.4.3. Demographic and Baseline Characteristics 33
6.4.4. General Medical History and Disease-Specific Medical History 34
6.4.5. Prior and Concomitant Medication 34
6.5. Efficacy Analysis 34
6.5.1. General Analysis Methods 34
6.5.2. Analysis of Primary Efficacy Endpoint(s) 35
6.5.2.1. Sensitivity Analysis to Primary Efficacy Endpoint(s) 36
6.5.2.2. Supplementary Analysis to Primary Efficacy Endpoint(s) 38
6.5.2.3. Subgroup analyses 40
6.5.3. Analysis of Secondary Efficacy Endpoint(s) 41
6.5.3.1. Analysis of Key Secondary Efficacy Endpoints 41
6.5.3.2. Analysis of Secondary Efficacy Endpoints Not Adjusted for Multiplicity 44
6.5.3.3. Sensitivity Analyses for Secondary Efficacy Endpoints 46
6.5.3.4. Subgroup analyses 48
6.5.3.5. Supplementary analyses 48
6.5.4. Analysis of Other Efficacy Endpoint(s) 48
6.5.4.1. Supplementary analyses 51
6.6. Safety Analysis 52
6.6.1. Adverse Events 52
6.6.1.1. Important Potential Risks and Adverse Events of Special Interest (AESI) 53
6.6.1.2. Potential Risk and Adverse Event of Special Interest (AESI) 54

6.6.1.3. Potential Risk 54
6.6.1.4. AAV Gene Therapy Class Effects 54
6.6.2. Clinical Laboratory Parameters 55
6.6.3. Vital Signs 56
6.6.4. Electrocardiogram 57
6.6.5. Liver Imaging 57
6.6.6. Other Safety Parameters 57
6.6.7. COVID-19 Impact on Subject Participation 57
7. CHANGES TO ANALYSIS SPECIFIED IN PROTOCOL 58
8. LIST OF PLANNED TABLES AND FIGURES 59
9. LITERATURE REFERENCES 60
10. APPENDICES 61
APPENDIX 1. HANDLING OF MISSING AND INCOMPLETE DATA 62
APPENDIX 2. VISIT WINDOWS 65
APPENDIX 3. CTCAE TOXICITY GRADES 67
APPENDIX 4. MEDDRA SEARCH STRATEGIES FOR IMPORTANT POTENTIAL RISKS, ADVERSE EVENTS OF SPECIAL INTEREST (AESI), POTENTIAL RISKS, AND AAV GENE THERAPY CLASS EFFECTS 68
APPENDIX 5. DERIVATION OF NON-INFERIORITY (NI) MARGINS 69
APPENDIX 6. DETERMINATION OF INTERCURRENT EVENTS 75

LIST OF TABLES

Table 1: Study Objectives and Endpoints 12
Table 2: Primary Estimand Attributes 15
Table 3: Key Secondary Estimand Attributes 16
Table 4: Prophylactic Oral Prednisolone/Placebo Taper Regimen 20
Table 5: Hochberg Method (Three Endpoints) 25
Table 6: Clinical laboratory Parameters 55
Table 7: Criteria for Clinically Significant Vital Signs 56

LIST OF FIGURES

Figure 1: Study Design Schematic, Study DTX401-CL301 21

Figure 2: Histogram of Difference in Mean Reduction in % Glucose Values <70 mg/dL Between DTX401 and Placebo 71

Figure 3: Histogram of Lower Limit of 95% Confidence Interval for Difference in Mean Reduction in % Glucose Values <70 mg/dL Between DTX401 and Placebo 71

Figure 4: Histogram of Difference in Mean Reduction in % Glucose Values 70-120 mg/dL Between DTX401 and Placebo 73

Figure 5: Histogram of Lower Limit of 95% Confidence Interval for Difference in Mean Reduction in % Glucose Values 70-120 mg/dL Between DTX401 and Placebo 74

TEAM PERSONNEL:

See eSignature Page

Signature

PPD

Print Name/Title

Date

SENIOR MANAGEMENT:

See eSignature Page

Signature

PPD

Print Name/Title

Date

See eSignature Page

Signature

PPD

Print Name/Title

Date

See eSignature Page

Signature

PPD

Print Name/Title

Date

LIST OF ABBREVIATIONS

Abbreviation	Definition
AAV	adeno-associated virus
AAV8	adeno-associated virus serotype 8
ACTH	adrenocorticotropic hormone
AE	adverse event
AESI	adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ARD	absolute relative difference
AST	aspartate aminotransferase
ATC	Anatomic Therapeutic Class
CDF	cumulative distribution function
CFC	controlled fasting challenge
CGM	Continuous Glucose Monitoring
CIC	Central Independent Committee
CMH	Cochran-Mantel-Haenszel
COVID-19	coronavirus disease 2019
CRF	case report form
CS	cornstarch
CTCAE	Common Terminology Criteria for Adverse Events
DRG	dorsal root ganglion
DMC	Data Monitoring Committee
DMP	Disease Monitoring Program
ECG	electrocardiogram
eDiary	Electronic diary
FUAP	Follow-up period
G6Pase	glucose-6-phosphatase (protein)
GC	genome copy
GEE	Generalized Estimating Equations
GGT	gamma-glutamyl transferase
GSD FAD	Glycogen Storage Disease Functional Assessment Diary
GSDIa	Glycogen Storage Disease Type Ia
HDL	high density lipoprotein
ICE	intercurrent event
ICH	International Conference on Harmonization
INR	international normalized ratio

Abbreviation	Definition
IP	investigational product
IRR	Infusion related reactions
IRT	Interactive Response Technology
ITT	Intention-to-Treat
IV	intravenous
LDL	low density lipoprotein
LFT	liver function tests
MAR	missing at random
MARD	mean absolute relative difference
MCMC	Markov-Chain Monte Carlo
mITT	Modified Intention-to-Treat
MNAR	missing not at random
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model for Repeated Measures
MRI	magnetic resonance imaging
MRU	medical resource utilization
PEAP	Primary Efficacy Analysis Period
PGIC	Patient Global Impression of Change
PGIC-CTT	Patient Global Impression of Change – Cornstarch Treatment Tolerance
PGIS	Patient Global Impression of Severity
PI	Principal Investigator
PP	Per Protocol
PT	preferred term
SAE	serious adverse event
SAP	Statistical Analysis Plan
SD	standard deviation
SE	standard error
SMBG	self-monitored blood glucose
SMT	Study Management Team
SMQ	Standardized MedDRA Query
Tab	total antibody
TEAE	treatment-emergent adverse event
TTH	time to hypoglycemia
ULN	upper limit of normal
WHO	World Health Organization

DOCUMENT REVISION HISTORY

Not applicable.

1. INTRODUCTION AND OBJECTIVES OF ANALYSIS

1.1. Introduction

Study DTX401-CL301 is a randomized, double blind, placebo-controlled Phase 3 clinical study of DTX401 in patients with glycogen storage disease type Ia (GSDIa). The purpose of the study is to evaluate the safety and efficacy of DTX401 in patients with glycogen storage disease type Ia (GSDIa) aged 8 years and above.

Patients with GSDIa have a deficiency in the glucose-6-phosphatase protein (G6Pase) enzyme and therefore cannot release glucose from glycogen storage in the liver during times of fasting (ie, during the period between meals) and are at risk for severe fasting hypoglycemia. DTX401 therapy is expected to provide G6Pase activity resulting in efficient release of glucose from glycogen stores in the liver, thereby reducing the risk of hypoglycemia during periods of fasting and the long-term complications associated with GSDIa.

1.2. Objectives of Analysis

The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within the DTX401-CL301 global Protocol (Amendment 3, version 4.0) dated 06 July 2023.

2. STUDY OBJECTIVE(S)

Table 1: Study Objectives and Endpoints

OBJECTIVES	ENDPOINTS
Primary	
To evaluate the efficacy of DTX401 to reduce or eliminate dependence on exogenous glucose replacement therapy needed to maintain glucose control	Percent change from Baseline to Week 48 in daily cornstarch intake for the DTX401 Group compared with the Placebo Group
Secondary (subject to multiplicity)	
To evaluate the effect of DTX401 on reducing the frequency of exogenous glucose replacement therapy	Change from Baseline to Week 48 in number of total daily doses of cornstarch
To evaluate the effect of DTX401 on glucose control	Change from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (<70 mg/dL [3.9 mmol/L]), for non-inferiority ; if non-inferiority is established, the endpoint will be tested for superiority
To evaluate the effect of DTX401 on subject experience of disease	PGIC assessment score at Week 48
Secondary (not subject to multiplicity)	
To evaluate the effect of DTX401 on glucose control	• Change from Baseline to Week 48 in time to hypoglycemia (<54 mg/dL [3.0 mmol/L]) during the CFC • Change from Baseline to Week 48 in percentage of glucose values in the range of 70-120 mg/dL [3.9-6.7 mmol/L], for non-inferiority ; if non-inferiority is established, the endpoint will be tested for superiority
To evaluate the safety of DTX401	Incidence, severity, and relationship to investigational product of TEAEs, TEAEs of special interest, serious TEAEs, related TEAEs, discontinuations from study or investigational product due to AEs, and fatal AEs

OBJECTIVES	ENDPOINTS
Tertiary	
To explore the effect of DTX401 on additional parameters of metabolic control	Change from Baseline to Week 48 (DTX401 vs placebo) in the following parameters: • Total number of hypoglycemic events (<54 mg/dL) in the 4 weeks prior to the visit based on CGM data. A hypoglycemic event is defined as at least 2 sensor values < 54 mg/dL that are at least 15 minutes apart with no intervening values >54 mg/dL. The end of the hypoglycemic event is defined as 2 sensor values > 70 mg/dL that are at least 15 minutes apart with no intervening values < 70 mg/dL. • Total duration of hypoglycemic events in the 4 weeks prior to the visit based on CGM data. • Time spent below 70 mg/dL (3.9 mmol/L) in the 4 weeks prior to the visit based on CGM data. • Observed overnight fasting time, based on PI assessment. • Pre-CFC whole blood lactate • Plasma uric acid • Pre-CFC plasma alanine • Daily dietary (non-cornstarch) macronutrient intake (fat, protein, carbohydrates, and non-utilizable sugars) and percentage of total calories from individual macronutrients • Time between a glucose level of 70 mg/dL (3.9 mmol/L) to < 54 mg/dL (3.0 mmol/L) during the CFC • Percentage of subjects who do not require overnight cornstarch (overnight cornstarch is defined as any dose disrupting normal sleep patterns) • Percentage of subjects who do not require cornstarch during the day • Percentage of subjects with > 75% reduction in total daily cornstarch intake

OBJECTIVES	ENDPOINTS
To explore the additional effect of DTX401 on subject experience of disease	• Change in GSD FAD symptoms score(s) from Baseline to Week 48 • Change in the GSD FAD impact score(s)from Baseline to Week 48 • Change in PGIS assessment score from Baseline to Week 48 • PGIC-CTT assessment score at Week 48 • Change in patient-reported health status via the EQ-5D-5L/ EQ-5D-Y instrument from Baseline to Week 48 • Change in MRU, specifically hospitalization and ER visits from Baseline to Week 48 • Change in work and school status and productivity, assessed via questionnaire and patient exit interviews from Baseline to Week 48 • Patient exit interviews
To explore the hormonal and metabolic responses to fasting in the CFCs	Characterization and change from baseline (as applicable) in the hormonal (ACTH, cortisol, IGFBP1, insulin, C-peptide) and metabolic (fatty acids, ketones, lactate, alanine) responses to fasting during the CFCs (DTX401 vs. placebo)
To explore additional safety of DTX401	Change from Baseline (DTX401 vs Placebo) in the following parameters: • eGFR, 24-hour urine protein, and urine protein-to-creatinine ratio • Hemoglobin, hematocrit, and platelets • Total bilirubin, ALT, AST, GGT, alkaline Phosphatase, albumin, INR • Heart rate, systolic and diastolic blood pressure, respiratory rate, weight (adults), weight Z-score (children), BMI (adults), BMI Z-score (children) • Triglycerides, Total cholesterol, LDL, and HDL • Presence and titer of anti-AAV8 antibodies • Presence and titer of anti-G6Pase antibodies • Cell-mediated response to AAV8 and G6Pase

AAV8, adeno-associated virus serotype 8; AE, adverse event; BMI, body mass index; CFC, controlled fasting challenge; CGM, continuous glucose monitoring; eGFR, estimated glomerular filtration rate; EQ-5D-(5L/Y), EuroQol- 5 Dimension (5-level/youth version); ER, emergency room; G6Pase, glucose-6-phosphatase; GSD FAD, Glycogen Storage Disease Functional Assessment Diary; GSDIa, glycogen storage disease type Ia; MRI, magnetic resonance imaging; MRU, medical resource utilization; PGIC, Patient Global Impression of Change; PGIC-CTT, Patient Global Impression of Change – Cornstarch Treatment Tolerance; PGIS, Patient Global Impression of Severity; TEAE, treatment-emergent adverse event.

2.1. Estimands

The primary estimand attributes are described in Table 2.

Table 2: Primary Estimand Attributes

Estimand Attribute	**Description**	
Population	Adults and pediatric ≥ 8 years of age with GSDIa	
Treatment	Subjects randomized to DTX401 vs. Placebo	
Endpoint	Percent change from Baseline to Week 48 in daily cornstarch intake	
Intercurrent Event 1	**Strategy**	**Description**
Transient changes in dietary management or nutritional intake due to acute events, such as trauma, surgery, infection, metabolic decompensation, etc.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.
Intercurrent Event 2	**Strategy**	**Description**
Persistent changes in dietary management or nutritional intake due to intentional or foreseeable physiologic changes, such as seasonal changes in activity, a new weight loss regimen, exercise program, puberty, pregnancy, etc.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.
Intercurrent Event 3	**Strategy**	**Description**
Sub-acute changes in glucose homeostasis due to adjustments in corticosteroid dose or other medications affecting glucose homeostasis, such as insulin.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.
Intercurrent Event 4	**Strategy**	**Description**
Partial dosing i.e., IP infusion stopped before completion and was not resumed due to adverse event (e.g., infusion related reaction) or technical issues.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.
Population-level summary	Difference in mean percent change from Baseline to Week 48 between two treatment groups	

Estimand attributes for key secondary efficacy endpoints are described in Table 3.

Table 3: Key Secondary Estimand Attributes

Endpoint 1: Change from Baseline to Week 48 in number of total daily doses of cornstarch		
Estimand Attribute	**Description**	
Population	Adults and pediatric ≥ 8 years of age with GSDIa	
Treatment	Subjects randomized to DTX401 vs. Placebo	
Endpoint	Change from Baseline to Week 48 in number of total daily doses of cornstarch	
Intercurrent Event 1	**Strategy**	**Description**
Transient changes in dietary management or nutritional intake due to acute events, such as trauma, surgery, infection, metabolic decompensation, etc.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.
Intercurrent Event 2	**Strategy**	**Description**
Persistent changes in dietary management or nutritional intake due to intentional or foreseeable physiologic changes, such as seasonal changes in activity, a new weight loss regimen, exercise program, puberty, pregnancy, etc.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.
Intercurrent Event 3	**Strategy**	**Description**
Sub-acute changes in glucose homeostasis due to adjustments in corticosteroid dose or other medications affecting glucose homeostasis, such as insulin.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.
Intercurrent Event 4	**Strategy**	**Description**
Partial dosing i.e., IP infusion stopped before completion and was not resumed due to adverse event (e.g., infusion related reaction) or technical issues.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.
Population-level summary	Difference in mean change from Baseline to Week 48 between two treatment groups	
Endpoint 2: Change from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (<70 mg/dL [3.9 mmol/L]), for non-inferiority		
Estimand Attribute	**Description**	
Population	Adults and pediatric ≥ 8 years of age with GSDIa	
Treatment	Subjects randomized to DTX401 vs. Placebo	
Endpoint	Change from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (<70 mg/dL [3.9 mmol/L])	
Intercurrent Event 1	**Strategy**	**Description**
Transient changes in dietary management or nutritional intake due to acute events, such as trauma,	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by

surgery, infection, metabolic decompensation, etc.		this intercurrent event will be used in the primary analysis.
Intercurrent Event 2	**Strategy**	**Description**
Persistent changes in dietary management or nutritional intake due to intentional or foreseeable physiologic changes, such as seasonal changes in activity, a new weight loss regimen, exercise program, puberty, pregnancy, etc.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.
Intercurrent Event 3	**Strategy**	**Description**
Sub-acute changes in glucose homeostasis due to adjustments in corticosteroid dose or other medications affecting glucose homeostasis, such as insulin.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.

Intercurrent Event 4	**Strategy**	**Description**
Partial dosing i.e., IP infusion stopped before completion and was not resumed due to adverse event (e.g., infusion related reaction) or technical issues.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.
Population-level summary	Difference in mean change from Baseline to Week 48 between two treatment groups	
Endpoint 3: PGIC assessment score at Week 48		
Estimand Attribute	**Description**	
Population	Adults and pediatric ≥ 8 years of age with GSDIa	
Treatment	Subjects randomized to DTX401 vs. Placebo	
Endpoint	PGIC assessment score at Week 48	
Intercurrent Event	**Strategy**	**Description**
Partial dosing i.e., IP infusion stopped before completion and was not resumed due to adverse event (e.g., infusion related reaction) or technical issues.	ITT Principle (Treatment Policy Strategy)	No adjustment will be made for the results affected by the intercurrent event. The assessments affected by this intercurrent event will be used in the primary analysis.
Population-level summary	Difference in mean score at Week 48 between two treatment groups	

3. STUDY DESIGN

Study DTX401-CL301 is a Phase 3, randomized, double-blind, placebo-controlled study to determine the efficacy and confirm the safety of DTX401 in subjects 8 years and older with GSDIa.

This study includes the following sequential stages:

Screening (Day -122 to Day -10): All screening evaluations must be completed by the Investigator and reviewed to confirm that subjects meet all eligibility criteria.

After first week of screening, the Investigator and/or study dietitian will review the subject's current diet, cornstarch regimen, and glycemic control to determine if the subject requires nutrition modifications or if the subject's diet, cornstarch, and glycemic control are considered optimized. If a subject does not require nutrition optimization, they may begin the 4-week stabilization period immediately.

- Nutrition Optimization Period: The goal of nutrition optimization is to adjust cornstarch dosing and diet to minimize the occurrence of hypoglycemia and minimize elevations in lactate, triglycerides, and uric acid, while avoiding excessive cornstarch consumption. Once the Investigator and/or dietitian determine the subject has achieved nutritional optimization, the subject will begin the nutrition stabilization period.
- Nutrition Stabilization Period: All subjects must demonstrate at least 4 consecutive weeks of glycemic, diet, and cornstarch stability prior to randomization.

Randomization Period (Day -10 to Day 0): Using the interactive response technology (IRT) system, the site will randomly assign the subject to DTX401 or placebo treatment and schedule the investigational product (IP, i.e., DTX401 or placebo) infusion at least 10 business days before the IP infusion date to allow time for IP and the prophylactic oral prednisolone/placebo kit to be delivered to the site.

Baseline (Day 0): The day before IP administration (i.e., Day 0), the subject will be admitted to the hospital or a research facility for baseline assessments.

Day 1 IP Administration: On Day 1, subjects randomly assigned to DTX401 will receive a single, blinded, peripheral IV infusion of DTX401 at 1.0×10^{13} GC/kg. Subjects randomly assigned to placebo will receive a single, blinded, peripheral IV infusion of normal saline.

Day 15: Subjects randomly assigned to DTX401 will start an oral prophylactic prednisolone regimen to minimize or prevent potential vector-induced hepatic effects (e.g., transaminase elevations). Subjects randomly assigned to placebo IP infusion will receive a matching oral placebo regimen to maintain the study blind. Oral prednisolone (or placebo) will be sustained for 4 weeks, and then tapered over approximately 4 weeks as shown in Table 4.

Table 4: Prophylactic Oral Prednisolone/Placebo Taper Regimen

Prednisolone/Placebo Dose			Monitoring Phone Calls
Duration of Dose	**Subjects Weighing ≥ 60 kg**	**Subjects Weighing < 60 kg[a]**	
4 weeks	60 mg/day[b]	1.0 mg/kg/day[b]	Investigators will call subjects to monitor for symptoms or signs of iatrogenic Cushing's syndrome and adrenal insufficiency at the end of each taper stage
5 days	40 mg/day	0.7 mg/kg/day	
5 days	30 mg/day	0.5 mg/kg/day	
5 days	20 mg/day	0.3 mg/kg/day	
5 days	10 mg/day	0.2 mg/kg/day	
10 days	5 mg/day	0.1 mg/kg/day[c]	

[a] Doses for subjects weighing < 60 kg should be rounded down to the nearest 5 mg .

[b] Prednisolone dose for subjects experiencing increased lactate or intolerable steroid toxicity after the initial 2 weeks of treatment will be reduced to 50 mg/day or 0.8 mg/kg (as applicable). If the intolerable effects do not begin to improve within 5 days of the dose adjustment, the subject may begin tapering (to 40 mg/day or 0.7 mg/kg/day, as applicable).

[c] For the last 10 days of dosing for subjects weighing less than 25 kg, the dose should be 5 mg every other day instead of rounding the dose down to 0.

The Primary Efficacy Analysis Period (PEAP; Day 1 to Week 48): The PEAP is the time from Day 1 post-dose through the completion of the Week 48 CFC.

Week 48 IP Administration: At Week 48, after all Week 48 assessments are completed, eligible subjects will receive a second blinded infusion of IP: subjects randomly assigned to DTX401 at baseline will receive placebo at Week 48, while eligible subjects randomly assigned to placebo at baseline will receive a single, blinded, peripheral IV infusion of DTX401 at 1.0×10^{13} GC/kg at Week 48.

It is possible that a subject who received placebo on Day 1 may not be able to receive DTX401 at Week 48 due to development of anti-AAV8 antibodies or due to another postbaseline change in disease status or overall health. Subjects will begin an oral prophylactic prednisolone regimen or matching oral placebo regimen 14 days after the Week 48 IP administration.

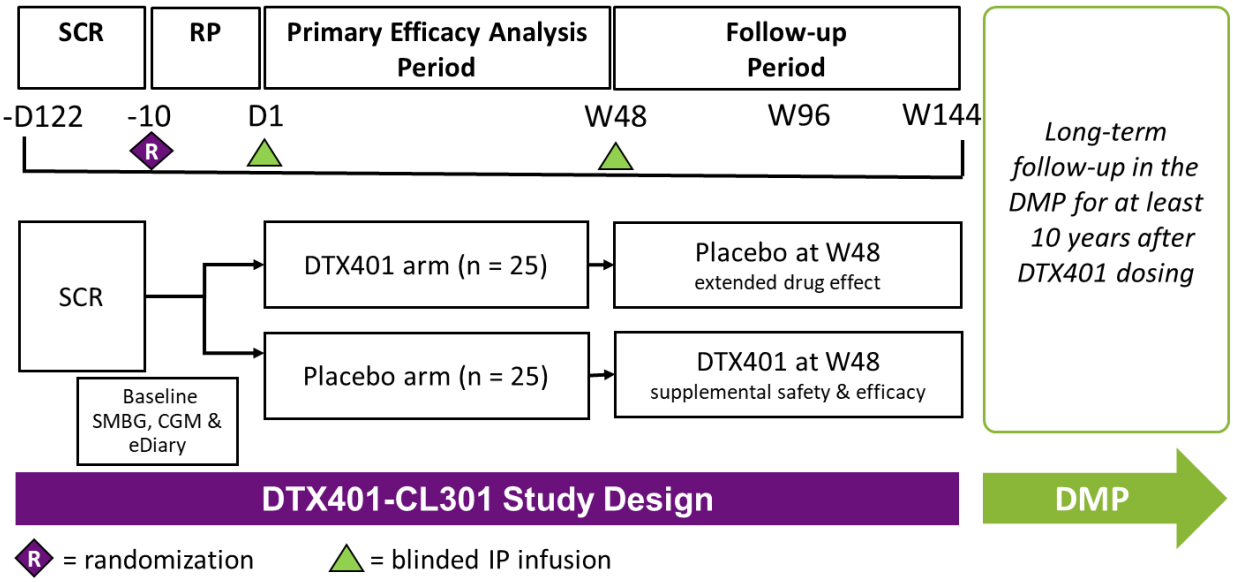
Holding Period (if required): The holding period consists of clinical, or home health visits every 12 weeks as needed to allow continued follow-up of subjects following the PEAP (e.g., due to interruption to IP supply). If proceeding with the second IP administration, the timing of all subsequent visits and assessments should be anchored to the date of the second IP administration.

The Follow-up Period (FUAP: Week 48 through Week 144): Study assessments during the FUAP will further characterize the efficacy and safety of DTX401 administration (up to 144 weeks after DTX401 treatment) to inform the overall risk-benefit profile.

Upon completion of this study (Week 144 or Early Withdrawal), all subjects who receive DTX401 will be asked to enroll in a long-term follow-up study (GSDIa DMP) to evaluate the safety and effectiveness of DTX401 via the DMP for at least 10 years after DTX401 administration.

A study design schematic for Study DTX401-CL301 is presented in Figure 1.

Figure 1: Study Design Schematic, Study DTX401-CL301



Abbreviations: CGM, continuous glucose monitoring; D, day; DMP, disease monitoring program; eDiary, electronic diary (for tracking cornstarch and dietary intake, and self-measured capillary blood glucose measurements [SMBG]); IP, investigational product (DTX401 or Placebo); W, week.

3.1. Study Population

The patient population for this study includes approximately 50 GSDIa patients who satisfy eligibility criteria mentioned in the protocol. Approximately 20 subjects < 18 years of age will be enrolled. Key eligibility criteria are described below.

Key Inclusion Criteria:

- Male and female patients ≥ 8 years of age at the time of informed consent.
- Subject has a diagnosis of GSDIa confirmed by deficient enzymatic activity (on liver biopsy), or by molecular testing of *G6PC* gene revealing 2 pathogenic mutations; in case of a single pathogenic mutation: clinical diagnosis is compatible with GSDIa and absence of characteristic features of GSDIb (i.e., chronic neutropenia, inflammatory bowel disease).
- Subject is currently receiving a therapeutic regimen of cornstarch (or equivalent), in accordance with nutrition guidelines (detailed in the Study Reference Manual) with stable nutrition, glycemic, and clinical status, as evidenced by:
  - No more than a 10% variation in weekly average daily cornstarch (or equivalent) intake over the last 4 weeks.
  - No more than a 25% variation in weekly average daily non-cornstarch carbohydrate over the last 4 weeks.

  - No more than 15% variation in weekly percentage of values in the target blood glucose range (60-120 mg/dL) over the last 4 weeks as measured by CGM and corroborated by SMBG. If adequate corroboration is not observed, this assessment should be made by SMBG.
  - No hospitalization for hypoglycemia and no severe hypoglycemic event (SHE) during the 4-week period preceding randomization and dosing, notwithstanding events of hypoglycemia due to unavoidable and unforeseeable events (e.g., infection, trauma) that transiently prevent the subject from tolerating enteral intake or acutely change the subject's metabolic demands, provided that the subject quickly returns to their prior physiologic state.

Key Exclusion Criteria:

- Detectable pre-existing antibodies to the AAV8 capsid
- History of liver transplant, including hepatocyte cell therapy/transplant
- History of hepatic fibrosis or cirrhosis
- Presence of liver adenoma > 5 cm in size or presence of liver adenoma > 3 cm and ≤ 5 cm in size with a documented annual growth rate of ≥ 0.5 cm per year
- Significant hepatic injury or dysfunction as evidenced by imaging or any of the following laboratory abnormalities from 2 consecutive samples (collected at least 4 weeks apart):
  - Alanine aminotransferase (ALT) or aspartate aminotransferase (AST) > 2.5 × the upper limit of normal (ULN)
  - Total bilirubin > ULN unless the subject has Gilbert syndrome
  - Alkaline phosphatase > ULN, with gamma-glutamyl transferase (GGT) > ULN
- Non-fasting triglycerides ≥ 1000 mg/dL

3.2. Dosage and Administration

Subjects meeting the eligibility criteria will be randomly assigned in a 1:1 ratio to receive DTX401 (n = 25) or Placebo (n = 25). On Day 1, subjects assigned to DTX401 will receive a single, blinded, peripheral IV infusion of DTX401 at 1.0×10^{13} gene copies (GC)/kg. Subjects assigned to placebo will receive a single, blinded, peripheral IV infusion of normal saline. At Week 48, eligible participants will cross over and receive DTX401 if they had previously received placebo or placebo if they had previously received DTX401 and will be followed for an additional 96 weeks.

The planned study duration is approximately 162 weeks (with possible extension if a subject enters a holding period), including the following:

- Pre-dosing Period of up to 18 weeks
- PEAP of 48 weeks
- Follow-up Period of 96 weeks

3.3. Blinding and Randomization Methods

3.3.1. Blinding Strategy

Study DTX401-CL301 uses a double-blind placebo-controlled design. Both IP (DTX401 or placebo) and the oral prednisolone/placebo taper will be blinded. The blind will be maintained for the subject, Investigator, and sites, to the extent that is safe and medically feasible, during the study until the last subject completes Week 96 visit or withdraws early. Site-facing Sponsor personnel will remain blinded until the last subject completes Week 96 visit or withdraws early. Some Sponsor personnel who are not site-facing will become unblinded to perform the PEAP analysis following database lock for the PEAP.

Baseline characteristics and assessments are not blinded in order to confirm that subjects comply with eligibility criteria and for awareness of the baseline characteristics of the study population.

Further details regarding the blinding strategy are provided in the Study Protocol.

3.3.2. Randomization

Using the IRT system, the site will randomly assign the subject to DTX401 or placebo treatment and schedule the investigational product infusion at least 10 business days before the IP infusion to allow time for IP and the prophylactic oral prednisolone/placebo kit to be delivered to the site.

3.4. Potential Covariate(s)

The primary and secondary analyses based on the use of the MMRM will consider baseline value; stratification factors (categories of age and baseline percentage of values within the 60 to 120 mg/dL glucose range as captured in the IRT), the treatment group, study visit (Weeks 24, 30, 36, 42 and 48), and interaction of visit by treatment as fixed effects.

3.5. Subgroups

Subgroup analyses will be performed for the primary and secondary endpoints by age (< 18 years vs ≥ 18 years) and baseline percentage of values in the glucose range of 60 to 120 mg/dL ($< 60\%$, $\geq 60\%$ and $\leq 85\%$, $> 85\%$) as captured in the IRT.

Subgroup analysis may be performed for pediatric subjects for pubertal stage status at baseline as per the Tanner pubertal stage development scale.

Subgroup analysis will also be performed by geographical region (e.g., North America, Europe, South America, Asia, etc.).

Subgroup analysis for primary and key secondary efficacy data and key safety data may be performed for subjects who had detectable AAV8 antibodies at baseline visit (Day 0 visit for subjects treated in PEAP and Week 48 visit for subjects treated in FUAP).

3.6. Sample Size Considerations

For the primary endpoint, a sample size of 50 subjects (25 subjects in the DTX401 Group, 25 subjects in the Placebo Group) will provide approximately 91% power to detect a difference between the DTX401 Group and the Placebo Group at Week 48 using MMRM including repeated measures at Week 24, Week 30, Week 36, Week 42, and Week 48 at 0.05 alpha. This

assumes a minimum average cornstarch reduction of 50% or more in the DTX401 Group at Week 48 versus a maximum average of 10% cornstarch reduction or less in the Placebo Group at Week 48 with a common SD of 40%, a correlation coefficient of 0.4 between assessments, and a discontinuation rate of 10% at Week 48.

The assumptions included in the sample size calculations are based on the results observed in an ongoing DTX401 Phase 1/2 (401GSDIA01) interventional study. In Phase 1/2 study, the mean percent change from baseline in cornstarch use at Week 52 was ~70% (SD=23%), Week 52 completion rate of 92% and correlation coefficient of ~0.7 across visits. Conservative estimates of these observations were used as assumptions in the sample size calculation considering that Phase 3 is a blinded study and includes placebo group.

Ultragenyx proprietary SAS code was used to determine the power for various sample sizes under these assumptions. The code simulated 1000 datasets of fixed size under the assumptions using PROC SIMNORMAL and MMRM analysis was performed on each dataset. Power was estimated as the percentage of times a significant result at 0.05 alpha was obtained. The sample size providing at least 90% power for the primary endpoint was determined.

3.7. Statistical Hypothesis

The primary endpoint is the percent change in daily cornstarch intake from Baseline to Week 48 comparing DTX401 with Placebo. The statistical hypothesis for this endpoint is:

H0: Mean percent change from Baseline to Week 48 in daily cornstarch intake in DTX401 Group is the same as mean percent change from Baseline to Week 48 in daily cornstarch intake in Placebo Group

H1: Mean percent change from Baseline to Week 48 in daily cornstarch intake in DTX401 Group is different than mean percent change from Baseline to Week 48 in Placebo Group

3.8. Multiple Comparisons/Multiplicity

The overall type I error rate for testing primary and key secondary endpoints will be controlled at the 2-sided 0.05 alpha level. If the primary endpoint is met, the following 3 key secondary endpoints will be tested using the Hochberg method based on the rank order of the 2-sided p-values as shown in Table 5.

- Change from Baseline to Week 48 in number of total daily doses of cornstarch
- Change from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (<70 mg/dL [3.9 mmol/L]) for non-inferiority
- PGIC assessment score at Week 48

Table 5: Hochberg Method (Three Endpoints)

	p-value	Signif.	p-value	Signif.	p-value	Signif.
Endpoint (1)	≤ 0.05	Yes	> 0.05	No	> 0.05	No
Endpoint (2)		Yes	≤ 0.025	Yes	> 0.025	No
Endpoint (3)		Yes		Yes	≤ 0.0167	Yes

Abbreviation: Signif = significant.
Note: Endpoints are ranked by p-values from (1) to (3) from the largest to the smallest.

Hochberg procedure is chosen because it is known to provide adequate overall alpha-control for independent endpoint tests or for positively correlated dependent tests with standard test statistics in some cases (e.g., the test statistics are jointly bivariate normal). Various simulation experiments for the general case (e.g., for more than two endpoints with unequal correlation structures) indicate that the Hochberg procedure usually will, but is not guaranteed to, control the overall Type I error rate for positively correlated endpoints, but fails to do so for some negatively correlated endpoints (FDA guidance for industry "Multiple Endpoints in Clinical Trials" dated October 2022). Based on data from DTX401 Phase 1/2 (401GSDIA01) interventional study, we do not expect these endpoints to be highly positively or negatively correlated (i.e., |r| is expected to be <0.5 where r is correlation coefficient). If high negative correlations are observed in the study, Holm method will be used for multiplicity adjustment.

3.9. Interim Analysis

No interim analysis is planned for this study.

3.10. Study Oversight Committee

Study oversight is provided by an unblinded, independent DMC and an unblinded CIC.

3.10.1. Data Monitoring Committee

An independent DMC including 2 clinicians and 1 statistician will be constituted for Study DTX401-CL301 and will act in an advisory capacity to provide safety oversight and to monitor subject safety on a routine basis throughout the course of the study.

The DMC will meet and review study data at the frequency defined and in accordance with the scope and objectives set forth in the DMC Charter. The DMC may provide advice to Ultragenyx regarding the safety of subjects, the ethics of the study and the continuing scientific validity of the study. The DMC may also make recommendations to Ultragenyx concerning continuation, termination or other modifications of the study based on their review of data. Further details regarding the DMC will be included in the DMC charter.

The Ultragenyx Data Review Board will review the DMC recommendations and determine necessary actions that will be communicated accordingly to all stakeholders, e.g., Regulatory Authorities, IRBs/IECs, and Investigators.

3.10.2. Central Independent Committee

The CIC is composed of independent GSDIa experts who are unblinded to subjects' treatment assignments. The CIC will review CGM recordings, SMBG data, cornstarch use/dietary intake data on demand throughout the study to provide additional guidance to the Investigators as needed regarding cornstarch regimen modification. This guidance is to ensure that subjects will get the maximum benefit of their gene therapy by avoiding hyperinsulinemic responses driven by increased endogenous glucose production after DTX401 administration. The ultimate decision to reduce cornstarch intake and/or the rate of reduction will be each Investigator's responsibility, considering subjects' clinical status and overall safety. Additionally, the CIC will recommend an appropriate treatment regimen in response to transaminase elevations that occur after the prednisolone taper. CIC will review abnormal early morning cortisol and ACTH levels obtained 24 hours after completing the prednisolone/placebo taper.

4. DEFINITIONS

4.1. Analysis Study Periods

For analysis, PEAP is defined as the period prior to Week 48 dosing and FUAP is defined as the period after Week 48 dosing if subject receives dosing at Week 48 visit. Otherwise, PEAP is defined as period up to and including Week 48 visit and FUAP is defined as period after Week 48 visit.

4.2. Analysis Treatment Groups

Following treatment groups will be defined for analysis:

- DTX401: Subjects randomized to DTX401.
- Placebo: Subjects randomized to Placebo.
- Crossover DTX401: Placebo subjects who crossover and receive DTX401 at Week 48.
- Crossover Placebo: DTX401 subjects who crossover and receive Placebo at Week 48.
- Total DTX401: Subjects treated with DTX401 at any time during the study (Day 1 or Week 48).

4.3. Baseline

Unless specified otherwise, the following definitions will be used:

For DTX401 and Placebo group, baseline is defined as the last non-missing assessment prior to the first IP administration. For Crossover DTX401 group and Total DTX401 group, baseline is defined as last non-missing assessment prior to the DTX401 IP administration. Specifically for Crossover DTX401 group, the efficacy data at Week 48 visit will be considered as the baseline data for efficacy analyses because no efficacy data is expected to be collected after Week 48 assessments and prior to the dosing of DTX401 for this group.

4.4. Duration of exposure

The duration of exposure to IP (DTX401 or Placebo) during the PEAP is defined as follows for subjects receiving IP at Day 1:

Exposure (days) = end of PEAP – first IP dose date + 1

The duration of exposure to DTX401 for DTX401 and Crossover DTX401 group is defined as follows:

Exposure (days) = last date of follow-up in the study – DTX401 dose date + 1

Duration of exposure in days to prophylactic oral prednisolone/placebo medication during the PEAP is defined as:

Exposure (day) = last date of prednisolone/placebo administration in the PEAP – first dose date of prednisolone/placebo in the PEAP + 1 (if patients stop the medication and restart, the duration without medication will be deducted from the total exposure.)

For FUAP, Exposure (day) = last date of prednisolone/placebo administration in the FUAP – first dose date of prednisolone/placebo in FUAP + 1 (if patients stop the medication and restart, the duration without medication will be deducted from the total exposure.)

4.5. Study Day

If the visit date is on or after the date of first IP administration:

Study day = visit date – date of first IP administration + 1

If the visit date is prior to the date of first IP administration:

Study day = visit date – date of first IP administration

5. ANALYSIS SETS

5.1. Screened Set

The Screened Set will consist of all subjects who signed ICF and received a subject identification number.

5.2. Enrolled Set

Not applicable.

5.3. Full Analysis Set (FAS)

The full analysis set (FAS) consists of all randomized subjects.

5.4. Safety Analysis Set

The Safety Analysis Set consists of all subjects treated with study IP, including subjects who received partial dose of IP because infusion was stopped due to some reason (e.g., adverse event, technical issues) and not restarted. Subjects will be analyzed according to the treatment they received. Safety analyses will be performed on the Safety Analysis Set.

5.5. Intention to Treat Analysis Set (ITT)

The Intention-to-Treat (ITT) analysis set consists of all randomized subjects who received any IP (DTX401 or placebo), including subjects who received partial dose of IP. Analysis will be performed based on the randomized treatment in this study.

5.6. Modified Intention to Treat Analysis Set (mITT)

The modified Intention-to-Treat (mITT) consists of all randomized subjects who received any IP (DTX401 or placebo) including subjects receiving partial dose of IP and provided at least one post-dosing assessment pertaining to primary efficacy analysis (i.e., at least one assessment of daily cornstarch intake in 1 week period prior to any of the visits at Week 24, 30, 36, 40, 42, and 48). Analysis will be performed based on the randomized treatment in this study.

In this study, there is a minimum of 10 days between randomization and IP administration due to the time required to ship IP (DTX401/placebo) and prednisolone/placebo kits to the investigational site. Therefore, it is possible that a few subjects may not receive IP after being randomized (e.g., if the subject decides to withdraw from the study or due to some unforeseen circumstance).

All efficacy analyses will be performed on the mITT set. If all randomized subjects receive IP and provide at least 1 post-dosing assessment pertaining to primary efficacy analysis, all efficacy analyses will be performed on the ITT set as described above.

5.7. Per Protocol Set (PP)

The Per Protocol (PP) analysis set will exclude mITT patients with any major protocol deviations that may impact the primary and secondary efficacy analyses. The exclusion of patients from the PP analysis set will be determined in a final blinded data review meeting that will be held prior

to the disclosure of the randomization list. Reasons for exclusion will be documented for each patient. Supplemental analyses will be performed on the PP analysis set.

6. STATISTICAL ANALYSIS

The description of statistical methods in the SAP is definitive for all matters discussed in this section.

6.1. General Principles

Tabulations will be produced for appropriate demographic, baseline, efficacy, and safety parameters. Continuous variables will be summarized by number of subjects and mean, SD/SE, Q1, Q3, median, minimum, and maximum values. Categorical variables will be summarized by number and percentage of subjects.

All statistical tests will be two-sided and tested at a significance level of 0.05, unless stated otherwise. P-values will be rounded to and displayed in three decimals. If a p-value is less than 0.001, it will be shown in tables as <0.001. Confidence intervals will be 2-sided 95% confidence intervals, unless stated otherwise. There will be no adjustments for multiple comparisons, except that adjustment for multiplicity of testing will be performed on 1 primary and 3 key secondary efficacy endpoints with methods assuring the formal control of the Family Wise Error Rate (FWER) at a strict 5% level of significance. The multiplicity adjustment method to be used is fully described in Section 3.8.

Data will be mainly summarized by DTX401, Placebo, Crossover DTX401, and Total DTX401 group. IP administration and administration of Prednisolone/placebo will also be summarized for Crossover Placebo group.

All tabular summaries will be accompanied by individual subject listings. Handling of missing or incomplete data is described in Appendix 1. Analysis windows are described in Appendix 2.

The primary analysis for the study will be performed after the last subject completes the PEAP (Week 48) or an Early Withdrawal Visit. At this time, all analyses planned in this SAP will be performed using all data from the PEAP and available data from the FUAP at the time of primary database lock. Further analysis will be performed after the last study participant has completed the Week 96 visit; all analyses planned in this SAP will be performed at that time. The final database lock will occur after all subjects complete the FUAP or Early Withdrawal visit. At this time final analyses will be performed. All analyses planned in this SAP will be performed. No interim analyses are planned.

All output will be sorted and labeled according to the International Conference on Harmonization (ICH) recommendations and formatted to the appropriate page sizes. SAS® software version 9.4 or higher will be used to perform statistical analyses unless otherwise specified.

6.2. Subject Accountability

The number and percentage of subjects in analysis populations (ITT, mITT, PP, and safety) will be summarized. No formal statistical comparisons will be performed.

The number and percentage of subjects completing the study and discontinued from the study will be presented. The reasons for discontinuation will be summarized. Further, a separate disposition summary will be provided for PEAP and FUAP.

Screen-failure subjects (i.e., subjects screened but not randomized) and the associated reasons for failure to randomize will be tabulated overall for the Screened Population. The number and percentage of subjects who prematurely discontinue after randomization and prior to dosing will be presented. The reasons for discontinuation from study will be summarized.

The study duration (days) will be summarized for dosed subjects. The study duration is defined as data cut-off date or date of study completion/discontinuation (whichever comes first) – DTX401/placebo first dose date +1.

6.3. Protocol Deviations and Intercurrent Events

Major protocol deviations will be summarized by type for mITT set. Both major and minor protocol deviations will be listed.

At the discretion of the sponsor, major protocol deviations as determined by a review of the data prior to unblinding of the study results and the conduct of statistical analyses may result in the removal of a subject's data from the PP Analysis Set. The sponsor, or designee, will be responsible for producing the final protocol deviation file; this file will include a description of the protocol deviation, and clearly identify whether this deviation warrants exclusion from the PP Analysis Set. This file will be finalized prior to database unblinding.

Intercurrent events (ICEs) are the events that happen during the study and may affect the interpretation of the treatment effect or preclude the observation of the outcome of interest. Some of the events can also be protocol deviations. At the discretion of the sponsor (and after consulting with study investigational sites), ICEs impacting primary and secondary efficacy endpoints will be determined by a blinded review of the data prior to unblinding of the study treatment and conduct of PEAP analysis. Potential ICEs for the study and the process for determination of ICEs is described in Appendix 6. The sponsor will be responsible for producing the final ICE file; this file will include a description of the ICE, and clearly identify the subjects and their study visits impacted by that ICE. This file will be finalized prior to treatment unblinding. Listing of ICEs will be provided and sensitivity analyses will be performed to understand the impact of ICEs on primary and key secondary efficacy endpoints as described in Section 6.5.2.1.

6.4. Investigational Product Administration

6.4.1. Extent of Exposure

Dose of IP administered and exposure duration to IP will be summarized for the Safety analysis set. Additionally, information collected on Exposure eCRF related to infusion (e.g., volume prepared for infusion, staged infusion (Yes/No), total infusion time, infusion rate, total interruption time if infusion was interrupted, reason for interruption etc.) will be summarized.

Summary of dose administered, and infusion related information will be provided separately for dosing in PEAP (i.e., at Day 1) for DTX401 and Placebo group and dosing in FUAP (i.e., at Week 48) for Crossover DTX401 and Crossover Placebo group. Similar summary will also be provided for Total DTX401 group.

Duration of exposure to DTX401 and Placebo during PEAP will be provided. Further, duration of exposure to DTX401 during the study will be provided for Crossover DTX401 group, DTX401 group and Total DTX401 group.

Total dose and duration of exposure to prophylactic oral prednisolone (or placebo) will be summarized. Summary will be provided separately for prednisolone/placebo regimen during PEAP for DTX401 and Placebo group and prednisolone/placebo regimen during FUAP for Crossover DTX401 and Crossover Placebo group. Total dose and duration of exposure to prophylactic oral prednisolone will also be summarized for DTX401 group.

6.4.2. Measurement of Treatment Compliance

Subjects with significant non-compliance with IP dosing (e.g., received partial dose, did not get planned dose, received incorrect dose etc.) will be documented as major protocol deviation and will be summarized as part of major protocol deviations. Similarly, significant non-compliance with Prednisolone regimen will be documented as major protocol deviation and summarized as part of major protocol deviations.

Compliance with cornstarch and diet prescription

At each scheduled visit, investigator will compare actual cornstarch intake (including Glycosade) from subject eDiaries to prescribed cornstarch (including Glycosade) and if there is a clinically meaningful difference between prescribed and actual intake the investigator will document the action taken for mitigating these discrepancies (e.g., subject received nutrition counseling or prescription was modified and reason for modification of prescription) in "Prescribed Daily Cornstarch Intake" eCRF. The number and percentage of subjects with clinically meaningful difference between actual and prescribed cornstarch intake at each scheduled visit will be summarized along with the actions taken for mitigation for mITT set. Further, absolute relative difference between prescribed intake and actual intake at a visit will be summarized for subjects who received nutritional counseling as part of mitigation. Similar summary will also be provided for compliance between actual and prescribed non-cornstarch diet and absolute relative difference between prescribed carbohydrates intake and actual carbohydrates intake at a visit will be summarized for subjects who received nutritional counseling as part of mitigation.

6.4.3. Demographic and Baseline Characteristics

The following demographic parameters and baseline characteristics will be summarized descriptively for the ITT, mITT, PP, and Safety Analysis Set.

- age at consent (as captured on Demographics eCRF), age group, race, ethnicity, sex
- weight, height, and body mass index [BMI] at Baseline (Day 0) for adults and corresponding Z-scores and percentiles for pediatric subjects.
- randomization stratification factors.
- region, country

Listing showing randomization related information as captured in IRT (e.g., randomization number, treatment assignment, randomization stratum) will be provided.

Descriptive analyses and plots (if appropriate) will be provided to show optimization and stabilization of diet and cornstarch intake prior to randomization.

6.4.4. General Medical History and Disease-Specific Medical History

General medical history, GSDIa specific hepatic history, and GSDIa specific medical history related to diagnosis, glucose control and seizures, GSDIa management, and sleep disturbance will be summarized for the mITT set.

6.4.5. Prior and Concomitant Medication

Prior medication is defined as any medication started before the first IP dose date (medication start date prior to first IP dose date) for DTX401 and Placebo group.

Concomitant medication is defined as any medication taken on or after the first IP dose date [medication end date on or after first IP dose date (or ongoing)] for DTX401 and Placebo group. For Cross-over DTX401 group, concomitant medication is any medication taken on or after the DTX401 dose date. Both prior and concomitant medications will be coded by drug name and therapeutic class using WHO drug dictionary. Results will be tabulated by Anatomic Therapeutic Class (ATC) and preferred term for the Safety analysis set. If a subject took a specific medication multiple times or took multiple medications within a specific therapeutic class, that subject will be counted only once for the coded drug name or therapeutic class.

6.5. Efficacy Analysis

The efficacy analyses will be based on the mITT set.

6.5.1. General Analysis Methods

Mixed-Effects Model for Repeated Measures (MMRM):

For continuous efficacy endpoints, MMRM based on all observed data at Week 24 through Week 48 will be the primary analysis method to compare treatment effects.

Selection of MMRM as primary analysis method is based on Pharmaceutical Research and Manufacturers of America (PhRMA) recommendations in the published article "Recommendations for the Primary Analysis of Continuous Endpoints in Longitudinal Clinical Trials" which quotes "..Our main conclusion is that the mixed model approach is more efficient and reliable as a method of primary analysis, and should be preferred to the inherently biased and statistically invalid simple imputation approaches" (Mallinckrodt, 2008).

Further, several recent clinical studies in rare genetic diseases with continuous primary endpoints data have used MMRM as primary analysis method. e.g.,

- Libmeldy (gene therapy) for treatment of Metachromatic leukodystrophy (MLD) (2020)
- Xenpozyme (enzyme replacement therapy) for treatment of Acid Sphingomyelinase Deficiency (ASMD) (2022)

The MMRM will include fixed effect of treatment group, visit, and treatment group-by-visit interaction, the stratification factors and baseline measurement as covariates. Stratification

factors (age and percentage of values in target blood glucose range) will be included as categorical variables with 2 strata for age at screening (< 18 years vs ≥ 18 years) and 3 strata for the baseline percentage of values in the target blood glucose range (< 60%, 60- 85%, > 85%) using the strata captured in IRT at the time of randomization.

A compound symmetry covariance structure will be used to model the within-subject correlation over time. Considering the small size of the study, a parsimonious model (ie., smaller number of parameters to estimate) is considered more suitable for modelling correlation structure across time-points; hence, compound symmetry matrix is chosen. Further, data from DTX401 Phase 1/2 (401GSDIa01) study support this selection as follows:

1. Correlation between % reduction in daily cornstarch intake at Week 24 vs Week 36 and correlation between % reduction in daily cornstarch intake at Week 24 vs Week 52 based on 10 subjects were similar (0.71 and 0.77 respectively) and
2. SDs of % reduction in daily cornstarch intake at Week 24, 36 and 52 for 10 subjects were similar (21%, 23%, 24% respectively)

This shows that correlation is constant regardless of how far apart the measurements are, thus supporting the assumption of compound symmetry. If this model fails to converge, the alternative covariance structures will be tested. The covariance structure converging to the best fit, as determined by Akaike's information criterion, will be used as the main analysis.

MMRM estimates will be based on restricted maximum likelihood (REML) repeated measures approach. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom.

Although MMRM includes data from multiple visits in addition to Week 48, the hypothesis testing for efficacy endpoints will be based on treatment effect at Week 48. Specifically, the estimation of the Week 48 contrast (DTX401 − Placebo) will provide the Least Squares Mean, the 95% confidence interval of the estimated difference between DTX401 and Placebo, and the 2-sided p-value which will be used for the comparison between DTX401 and placebo.

6.5.2. Analysis of Primary Efficacy Endpoint(s)

The primary endpoint is the percent change in daily cornstarch intake from Baseline to Week 48 comparing DTX401 with Placebo.

The statistical hypothesis for this endpoint is:

H0: Mean percent change from Baseline to Week 48 in daily cornstarch intake in DTX401 Group is the same as mean percent change from Baseline to Week 48 in daily cornstarch intake in Placebo Group

H1: Mean percent change from Baseline to Week 48 in daily cornstarch intake in DTX401 Group is different than mean percent change from Baseline to Week 48 in daily cornstarch intake in Placebo Group

The endpoint will be evaluated using mixed model for repeated measures (MMRM). The model will include percent change from baseline in cornstarch intake at Weeks 24, 30, 36, 42, and 48.

For each post-baseline visit, the average of total daily cornstarch intake (including Glycosade) reported in the eDiary over 1 week prior to the visit will be calculated, except for Week 48 visit

where average over 4 weeks of data collected prior to the visit will be calculated, and the percent change from baseline will be included in the model. Baseline average cornstarch intake is calculated as the average of total daily cornstarch intake over 4 weeks prior to Baseline visit (Day 0).

The model will include the treatment group, and average cornstarch intake at Baseline, stratification factors, study visit (Weeks 24, 30, 36, 42, and 48), and interaction of visit by treatment as covariates. The estimation of the Week 48 contrast (DTX401 − Placebo) will provide the Least Squares Mean, the 95% confidence interval of the estimated difference between DTX401 and Placebo, and the 2-sided p-value for the comparison between DTX401 and placebo. If the treatment effect for DTX401 is greater than that for Placebo and the 2-sided p-value is less than 0.05, the superiority of DTX401 over Placebo will be demonstrated.

The descriptive summary statistics of daily cornstarch intake, change from baseline and percent change from baseline in daily cornstarch intake by study visit in each treatment group will be reported. Plots of mean percent change from baseline by study visits will be presented by treatment group.

As part of supplementary analyses, similar analyses will be presented for percent changes from baseline in daytime cornstarch intake, nighttime cornstarch intake, Glycosade intake, daytime Glycosade intake and nighttime Glycosade intake. Nighttime dose will be identified if subject responds “Yes” to the question “Did you wake up to take this dose?” in the eDiary for that dose; otherwise, dose will be considered as a daytime dose. MMRM will be used to estimate treatment difference if adequate number of subjects are available for the model, otherwise only descriptive statistics at each scheduled visit will be provided. All supplementary analyses based on MMRM will include nominal p-values.

6.5.2.1. Sensitivity Analysis to Primary Efficacy Endpoint(s)

Missing data: Sensitivity analysis will be performed by handling the missing data differently than considered in the main analysis for the primary endpoint. The strategy to handle the missing endpoint data in the main analysis is based on ITT principle (treatment policy strategy) i.e., no imputation of missing data will be performed.

Subject is expected to enter cornstarch intake information daily in eDiary for 1 week (7 days) prior to the visit (except for Baseline and Week 48 visit where subject is expected to provide cornstarch intake for 4 weeks prior to the visit) and average of daily intake values will be the cornstarch intake at that visit. If data entry is missing on any day, data available on remaining days will be used to calculate the average. If no data is entered in the 1-week period (or 4-week period for Baseline and Week 48 visit), the cornstarch intake will be missing and MMRM will be performed with missing endpoint data as-is.

The MMRM model handles missing endpoint values under the assumption of missing at random (MAR), thus estimating the outcome that would occur at Week 48 conditional on values observed at previous time points. Although MAR is more restrictive assumption compared to missing not at random (MNAR) assumption; it seems to be a reasonable assumption for missing data in this study. Close monitoring of subjects is planned in the study. For subjects who experience poor efficacy or safety outcomes at any visit, significant efforts will be made to retain them in the study and to collect required data at future visits. Thus, it is less likely that a subject

may drop-out without showing any signs of poor efficacy or safety at earlier visits. Therefore, MNAR assumption (i.e., missing data cannot be reasonably predicted based on existing collected data) seems unlikely for missing data.

Sensitivity analysis 1 will include data from only those weeks where substantial data is provided, specifically where cornstarch intake is reported for at least 4 days out of 7 days (e.g., at Week 48 visit, only weeks where cornstarch intake is reported for at least 4 days will be included in the analysis). This 4/7 rule is commonly employed for dealing with missing data in daily diary data to determine whether a patient's summary score is created or set to missing (Griffiths et al., 2022). If there is no substantial data available, the endpoint at that visit will be missing and MMRM will handle missing endpoint values with MAR assumption. This analysis will assess robustness of results of main analysis which includes all available data, including data from visits with substantial missing data.

Sensitivity analysis 2 will impute missing endpoint data at Week 24, 30, 36, 42, and 48 using Multiple Imputation (MI) technique which imputes the data under the assumption of missing at random (MAR). SAS PROC MI will be used to create multiply imputed datasets (e.g., 30 datasets). MMRM will be performed on each multiply imputed dataset, and results (p-values) will be combined using PROC MIANALYZE.

Sensitivity analysis 3 will be the tipping point analysis which will impute missing endpoint data at Week 24, 30, 36, 42, and 48 under the assumption of missing not at random (MNAR). Steps included in tipping point analysis are as follows:

1. Impute missing data under MAR assumption using Multiple Imputation method (as described above). SAS PROC MI will be used to create multiply imputed datasets (e.g., 30 datasets).
2. Shift imputed values for DTX401 and placebo by +/- delta using the MNAR statement in SAS PROC MI, perform MMRM on each dataset, and combine results (p-value) from these analyses using SAS PROC MIANALYZE
3. Repeat 1 and 2 above with varying delta till loss of significance of results is observed. These delta values are tipping points/scenarios.
4. Evaluate the clinical plausibility of tipping scenarios.

Intercurrent events (ICEs): ICEs are the events that affect the interpretation of the treatment effect or preclude the observation of the outcome of interest. The potential intercurrent events for primary endpoint are described in Table 2. Sensitivity analysis will be performed by handling the data impacted by ICEs differently than handled in the main analysis for the primary endpoint. The strategy to handle the ICEs in the main analysis for primary endpoint will be based on ITT principle (i.e., treatment policy strategy) i.e., the occurrence of the intercurrent event is considered irrelevant in defining the treatment effect of interest; no adjustment will be made to the endpoint values impacted by ICEs and MMRM will be performed with data as-is.

Sensitivity analysis will include hypothetical strategy to handle ICEs i.e., setting endpoint data impacted by ICEs to missing and imputing missing endpoint values using Multiple Imputation method. SAS PROC MI will be used to create multiply imputed datasets (e.g., 100 datasets) using Markov-Chain Monte Carlo (MCMC) and complete datasets will be analyzed with

MMRM model. Results (p-values) from the analyses will be combined using PROC MIANALYZE.

Normality assumption: Main analysis of the primary endpoint is based on the normality assumption for the primary endpoint because based on the data observed in the Phase 1/2 study (401GSDIA01), data are expected to follow normality. If a departure from normality is observed based on statistical tests and graphical methods, sensitivity analysis will be performed to assess the robustness of main analysis results to normality assumption. First, appropriate transformations of data will be performed using Box-Cox transformation and MMRM will be performed based on transformed data. If transformation does not result in normality, the weighted Generalized Estimating Equations (GEE) approach or Generalized Linear Mixed Models approach may be used.

Change in cornstarch collection from 4 weeks to 1 week prior to the visit during the study: As per schedule of events included in the original Protocol, Protocol Amendment 1 and Protocol Amendment 2, subjects were required to complete cornstarch entries daily for 4 weeks prior to the scheduled visit. That requirement was changed later during the study in the Protocol Amendment 3 to ease data entry burden on subjects and it required subjects to complete cornstarch entries daily for only 1 week prior to the visit, except for Baseline and Week 48 visit where the requirement to complete cornstarch entries for 4 weeks prior to the visit was not changed. Hence, it is likely that for a subject, cornstarch intake may have been collected over 4 weeks at some earlier visits in the study and for only 1 week at some later visits in the study.

To maintain consistency in the use of cornstarch data across all subjects in the analysis, only data collected during 1 week prior to the visit will be used (except for Baseline and Week 48 visit where 4 weeks of data prior to the visit will be used) in the main analysis of primary efficacy endpoint i.e., average of cornstarch intake over 7 days prior to the visit will be used (except for Baseline and Week 48 visit where average over 28 days prior to the visit will be used) as the cornstarch intake at a visit.

As a sensitivity analysis, all available data in 4 weeks prior to the visit will be used i.e., average of cornstarch intake available over 28 days prior to the visit will be used as the cornstarch intake at a visit. This analysis will assess robustness of results of main analysis to changes in cornstarch data collection strategy during the study.

Using median to calculate average cornstarch intake:

For the main/primary analysis, mean of daily cornstarch intake in 4 weeks or 1-week period prior to the visit is used to calculate average cornstarch intake at a visit. In case of outlier(s), median is a more robust measure compared to mean. Hence, as a sensitivity analysis, the median of daily cornstarch intake in 4 weeks or 1-week period is used to calculate average cornstarch intake at a visit and is further used to calculate the percent change from baseline in cornstarch intake at a visit. This analysis will assess robustness of results of main analysis to any outliers in cornstarch intake.

6.5.2.2. Supplementary Analysis to Primary Efficacy Endpoint(s)

Using prescribed cornstarch intake: The primary efficacy analysis will be repeated using the prescribed cornstarch intake documented in CRFs to understand the impact of subject compliance to the prescription of cornstarch on primary efficacy result. Plot showing mean

percent change in prescribed cornstarch intake along with mean percent change in actual cornstarch intake (captured in eDiary) by study visit will be presented by treatment group.

Using PP set: The primary efficacy analysis will be repeated on the PP set to evaluate the impact of important protocol deviations on the primary efficacy result.

Using ITT set: The primary efficacy analysis will be repeated on the ITT set to evaluate the impact of including subjects who were randomized and dosed but are not part of mITT set (since they did not provide post-dosing efficacy data pertaining to primary analysis) on the primary efficacy result. MMRM will handle the missing data for these subjects with MAR assumption.

Using FAS: The primary efficacy analysis will be repeated on the FAS (all randomized subjects) to evaluate the impact of including subjects who were randomized but were not part of mITT set (since they discontinued prior to the dosing, or they did not provide post-dosing efficacy data pertaining to primary analysis) on the primary efficacy result. MMRM will handle the missing data for these subjects with MAR assumption.

Using only Week 48 data: Analysis based on only baseline and Week 48 data will be performed as a supplementary analysis. Percent change from baseline in daily cornstarch intake at Week 48 will be compared between DTX401 and placebo group using linear model which will include fixed effect of treatment group, and stratification factors and baseline cornstarch intake as covariates. If there are missing data at Week 48, the analysis will be conducted in 2 different ways:

1. Using only available data at Week 48 (i.e., completers analysis).
2. Imputing missing data at Week 48 using multiple imputation method. The same linear model described above for the analysis will be used for multiple imputation.

Excluding subjects for whom cornstarch was not optimized at baseline: At baseline, subjects' cornstarch and diet are optimized to maintain glycemic control. If there are cases where subjects were dosed without optimization of cornstarch, those will be recorded as major protocol deviations and supplementary analysis will be performed where MMRM will be performed excluding these subjects.

Nominal p-values will be provided as part of the supplementary analyses described above.

Counts and percentages of subjects in following subcategories of cornstarch reduction from baseline will be reported by study visit and treatment group:

$< 50\%$, $\geq 50\%$, $\geq 70\%$, $\geq 90\%$, and 100%

Responder analysis:

Clinical meaningfulness of the primary endpoint will be supported with a responder analysis based on a $\geq 50\%$ reduction in cornstarch quantity (i.e., subject with daily cornstarch reduction $\geq 50\%$ from baseline to Week 48 will be considered as a responder). Responder rates will be compared using Cochran-Mantel Haenszel (CMH) test using stratification factors. An estimate of the difference in responder rates and corresponding 95% confidence interval (based on Wilson score method) will be provided. Nominal p-value will be provided. Additionally, cumulative distribution function (CDF) plots of responses for the DTX401 and placebo group showing percent change from baseline to Week 48 in daily cornstarch on the X-axis and the percent of patients experiencing that change on the Y-axis will be provided.

As part of Week 96 analysis and Final analysis, the following analyses will be performed:

- Percent change in cornstarch intake from baseline to 48 weeks after DTX401 treatment will be compared between DTX401 group and Crossover DTX401 group using the MMRM model. For DTX401 group, data at Week 24, 30, 36, 42, and 48 will be included in the model whereas for the Crossover DTX401 group, data at Weeks 72, 78, 84, 90, and 96 will be included in the model. These visits will be renamed as Weeks 24, 30, 36, 42, and 48 post-DTX401 treatment respectively and included in the model. The model will include the treatment group (DTX401 and Crossover DTX401), average cornstarch intake at Baseline, stratification factors, study visit, and interaction of visit by treatment as covariates. To calculate the change from baseline, baseline will be defined as Week 48 visit for Crossover DTX401 group and Day 0 visit for DTX401 group. The estimation of the Week 48 contrast (DTX401 − Crossover DTX401) will provide the Least Squares Mean, the 95% confidence interval of the estimated difference between DTX401 and Crossover DTX401 and the p-value for the comparison between DTX401 and Crossover DTX401.
- For the Crossover DTX401 group, the percent change in cornstarch intake before they received DTX401 treatment will be compared to that after they received DTX401 treatment. Specifically, the percent change from baseline to Week 48 (‘Before Treatment” period) will be compared with percent change from baseline to Week 96 (“After Treatment” period) for them using MMRM model. Their data at Weeks 24, 30, 36, 42, and 48 will correspond to “Before Treatment” period; data at Weeks 72, 78, 84, 90, and 96 will correspond to “After Treatment” period and will be included in the model as Weeks 24, 30, 36, 42, and 48 visit respectively. The model will include the period (Before Treatment and After Treatment), average cornstarch intake at Baseline, stratification factors, study visit, and interaction of visit by period as covariates. To calculate the change from baseline, baseline will be defined as Day 0 visit for "Before Treatment” period and Week 48 visit for "After Treatment” period. The estimation of the Week 48 contrast (Before Treatment− After Treatment) will provide the Least Squares Mean, the 95% confidence interval of the estimated difference between 2 periods and the p-value for the comparison between 2 periods.

6.5.2.3. Subgroup analyses

Subgroup analysis will be performed for the subgroups identified in Section 3.5. Further, subgroup analysis will be performed for following subgroups:

- subjects dosed with <4 and ≥4 average number of total daily doses of cornstarch (including Glycosade) at baseline.
- subjects using Glycosade (Yes/No) at baseline

Descriptive statistics of primary endpoint will be provided within each subgroup. Further, an estimate of the treatment difference based on the MMRM model with corresponding 95% CI will be presented within each subgroup for exploratory purposes. If subgroup size is very small, MMRM estimates will not be provided, and only descriptive statistics will be provided. Forest plots may be provided for subgroups.

6.5.3. Analysis of Secondary Efficacy Endpoint(s)

The MMRM method will be used for the analysis of the secondary endpoints for continuous outcomes.

6.5.3.1. Analysis of Key Secondary Efficacy Endpoints

Secondary efficacy endpoints included in this section will be adjusted for multiplicity using Hochberg procedure mentioned in Section 3.8.

6.5.3.1.1. Change from Baseline to Week 48 in Number of Total Daily Doses (Frequency) of Cornstarch

The statistical hypothesis for this endpoint is:

H0: Mean change from Baseline to Week 48 in number of total daily doses (frequency) of cornstarch in DTX401 Group is same as mean change from Baseline to Week 48 in total daily doses of cornstarch in Placebo Group

H1: Mean change from Baseline to Week 48 in number of total daily doses (frequency) of cornstarch in DTX401 Group is different than mean change from Baseline to Week 48 in number of total daily doses of cornstarch in Placebo Group

The endpoint will be evaluated using MMRM. MMRM Model will include change from baseline in number of total daily doses (frequency) of cornstarch intake (including Glycosade) at Weeks 24, 30, 36, 42, and 48. For each post-baseline visit, the average of daily number of doses of cornstarch intake over 1 week prior to the visit will be calculated, except for Week 48 visit where average over 4 weeks of data prior to the visit will be calculated and the change from baseline will be included in the analysis. The baseline total number of daily doses of cornstarch will be calculated as the average of daily number of doses of cornstarch intake over 4 weeks prior to the Baseline visit (Day 0).

The model will include the treatment group, and total daily doses of cornstarch intake at Baseline, stratification factors as captured in IRT, study visits, and interaction of visit by treatment as covariates. The estimation of the Week 48 contrast (DTX401 – Placebo) will provide the Least Squares Mean, the 95% confidence interval of the estimated difference between DTX401 and Placebo and the 2-sided p-value for the comparison between DTX401 and placebo. The p-value will be used in Hochberg procedure for ranking the endpoints and testing hypothesis.

The descriptive summary statistics of total daily cornstarch doses and changes from in each treatment group will be reported by scheduled visit. Plots of mean total daily cornstarch doses by study visits will be presented along with mean percent reduction in daily cornstarch intake by treatment group.

Similar analyses will be performed for changes from baseline in number of total daytime doses (frequency) and number of total nighttime doses (frequency) of cornstarch (doses that disrupt the regular night sleep pattern of the subject), total doses (frequency) of Glycosade, total daytime doses (frequency) of Glycosade, and total nighttime doses (frequency) of Glycosade as a supplementary analysis. MMRM will be used to estimate treatment difference if adequate number of subjects are available for the model, otherwise only descriptive statistics at each

scheduled visit will be provided. All supplementary analyses based on MMRM will include nominal p-values. Plots of mean total daytime cornstarch doses and nighttime doses by study visits will be presented by treatment group.

Further, average duration of dosing interval and change from baseline will be analyzed separately for daytime and nighttime doses for each treatment group as a supplementary analysis. For each dose, duration of dosing interval is calculated as the time (in hours) from current dose to the next dose. MMRM will be used to estimate treatment difference if adequate number of subjects are available for the model, otherwise only descriptive statistics at each scheduled visit will be provided. All supplementary analyses based on MMRM will include nominal p-values.

6.5.3.1.2. Change from Baseline to Week 48 in Percentage of Glucose Values in Hypoglycemic Range (< 70 mg/dL)

The following non-inferiority hypothesis will be tested for this endpoint with non-inferiority margin of 10% where reduction is defined as baseline value – post-baseline value. If there is decrease in percentage of glucose values in hypoglycemic range from baseline, the reduction will be positive whereas if there is increase from baseline, the reduction will be negative.

H0: Mean reduction from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (< 70 mg/dL [3.9 mmol/L]) in DTX401 Group – mean reduction from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (< 70 mg/dL [3.9 mmol/L]) in Placebo Group ≤ -10%

H1: Mean reduction from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (< 70 mg/dL [3.9 mmol/L]) in DTX401 Group – mean reduction from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (<70 mg/dL [3.9 mmol/L]) in Placebo Group > -10%

The derivation of non-inferiority (NI) margin of 10% for testing this endpoint is included in Appendix 5.

MMRM Model to test the hypothesis will include change from baseline in percentage of glucose values in hypoglycemic range at Weeks 24, 30, 36, 42, and 48. For each visit, the percentage of glucose values in hypoglycemic range will be calculated and the change from baseline will be included in the analysis. Percentage of glucose values is calculated as the number of glucose values in hypoglycemic range divided by the total available number of glucose measurements for the visit. Baseline percentage of glucose values in hypoglycemic range is calculated as the percentage of glucose values in hypoglycemic range prior to the Baseline (Day 0) visit.

The model will include the treatment group, and Baseline percentage of glucose values in hypoglycemic range, stratification factors as captured in IRT, study visits, and interaction of visit by treatment as covariates.

The estimation of the Week 48 contrast (DTX401 – Placebo) will provide the Least Squares Mean estimate of the difference between DTX401 and Placebo. The 2-sided p-value for the non-inferiority test will be calculated and used in the Hochberg procedure for ranking the endpoints., To test the non-inferiority hypothesis, (1-α/2)100% 1-sided confidence interval for the difference in mean reduction from baseline in percentage of glucose values in hypoglycemic range between DTX401 and Placebo will be calculated where α is determined by the Hochberg procedure.

Statistical non-inferiority will be demonstrated if a lower bound of that confidence interval is greater than -10%.

Only if non-inferiority is established, the following hypothesis will be tested at 0.05 alpha to test for superiority. This hypothesis testing will not be part of multiplicity adjustment. The 2-sided p-value for the comparison and the 95% confidence interval for the estimated difference in mean change from baseline in percentage of glucose values in hypoglycemic range between DTX401 and Placebo will be provided based on the MMRM model described above.

H0: Mean change from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (< 70 mg/dL [3.9 mmol/L]) in DTX401 Group is the same as mean change from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (< 70 mg/dL [3.9 mmol/L]) in Placebo Group

H1: Mean change from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (<70 mg/dL [3.9 mmol/L]) in DTX401 Group is different than mean change from Baseline to Week 48 in percentage of glucose values in hypoglycemic range (<70 mg/dL [3.9 mmol/L]) in Placebo Group

Additionally, the percentage of glucose values in hypoglycemic range and the change from baseline at all scheduled time points will be summarized by the treatment group. Plots of mean percentage of glucose values in hypoglycemic range by study visits will be presented by treatment group.

Both continuous glucose monitoring (CGM) data and capillary glucose measurements (SMBG) will be independently used to analyze this endpoint. Two analyses will be performed:

(1) Using CGM data collected in 4 weeks prior to the visit and

(2) Using SMBG data collected in 1 week prior to the visit

The main analysis (i.e., MMRM model, multiplicity adjustment, hypothesis testing) for this endpoint will be based on CGM data (1), while analysis based on SMBG data (2) will be considered as supplementary analysis and will be used to independently validate conclusions of main analysis.

Additional supplementary analysis will include MMRM analysis based on CGM data collected in 4 weeks prior to the visit but including CGM data from only those weeks where adequate corroboration between CGM and SMBG values (as described in Appendix 1) is observed.

All supplementary analyses will include nominal p-values.

6.5.3.1.3. PGIC Score at Week 48

The statistical hypothesis for this endpoint is:

H0: Mean PGIC score at Week 48 for DTX401 Group is same as mean PGIC score at Week 48 for Placebo Group

H1: Mean PGIC score at Week 48 for DTX401 Group is different than mean PGIC score at Week 48 for Placebo Group

PGIC score at Week 48 will be analyzed using MMRM. MMRM Model will include PGIC score at Weeks 24, 36, and 48. The model will include treatment group, stratification factors, study

visits, and interaction of visit by treatment as covariates. The estimation of the Week 48 contrast (DTX401 – Placebo) will provide the Least Squares Mean, the 95% confidence interval of the estimated difference between DTX401 and Placebo and the 2-sided p-value for the comparison between DTX401 and placebo. The p-value will be used in Hochberg procedure for ranking the endpoints and testing hypothesis.

Further, the PGIC score at all post-baseline scheduled time points will be summarized by treatment group. Plots of mean PGIC by post-baseline study visits will be presented by treatment group. The PGIC was administered during screening for the sole purpose of evaluating test-retest reliability of GSD FAD scores for GSD FAD psychometric validation. PGIC score at Screening will not be summarized but will be included in the listing.

6.5.3.2. Analysis of Secondary Efficacy Endpoints Not Adjusted for Multiplicity

Secondary efficacy endpoints included in this section will not be adjusted for multiplicity.

6.5.3.2.1. Change from Baseline to Week 48 in Time to Hypoglycemia During the CFC

The statistical hypothesis for this endpoint is:

H0: Mean change from Baseline to Week 48 in time to hypoglycemia during CFC in DTX401 Group is same as mean change from Baseline to Week 48 in time to hypoglycemia during CFC in Placebo Group

H1: Mean change from Baseline to Week 48 in time to hypoglycemia during CFC in DTX401 Group is different than mean change from Baseline to Week 48 in time to hypoglycemia during CFC in Placebo Group

MMRM Model will include change from baseline in time to hypoglycemia (TTH) during the CFC at Weeks 36, and 48. TTH (in hours) will be calculated based on the start and end times of CFC documented in the CRF and change from baseline will be included in the model. If subject does not experience hypoglycemia during the CFC, CFC will be stopped at 15 hours as per the protocol and that will be the end time of CFC noted in the eCRF. Baseline TTH is the TTH during the CFC performed on Day 0.

The model will include treatment group, and Baseline TTH, stratification factors, study visits, and interaction of visit by treatment as covariates. The estimation of the Week 48 contrast (DTX401 – Placebo) will provide the Least Squares Mean, the 95% confidence interval of the estimated difference between DTX401 and Placebo, and the 2-sided p-value for the comparison between DTX401 and placebo.

Additionally, the TTH and the change from baseline at all scheduled time points will be summarized by the treatment group. Plots of mean TTH by study visits will be presented by treatment group.

In this study, individualized amounts of cornstarch are being given to subjects prior to CFC which is intended to reflect their current cornstarch prescription. Since the amount of individualized cornstarch used for the CFC will vary across subjects, a supplementary analysis will be performed on normalized TTH by the amount of carbohydrates in cornstarch given prior to CFC using MMRM model which will include change from baseline in TTH/CS during the CFC at Weeks 36, and 48 where TTH/CS is TTH (in minutes) divided by amount of

carbohydrates (in grams) in cornstarch. The model will include treatment group, and Baseline TTH/CS, stratification factors as captured in IRT, study visits, and interaction of visit by treatment as covariates. TTH/ CS denotes the gain in TTH (in minutes) per gram of carbohydrates in cornstarch used for the CFC. If a subject does not receive cornstarch prior to the CFC, it will be included as 1 gram (very small quantity) for the calculation of TTH/CS. Further, if a subject takes cornstarch with any liquid other than water (e.g., a pediatric subject mixing cornstarch in a small dose of an infant formula), it is possible that it may impact TTH. Hence, the quantities of carbs (in grams) from the liquid will be added to the amount of carbohydrates in cornstarch and used in the analysis. It is expected that DTX401-treated subjects will reduce the amount of cornstarch needed to maintain normal glucose control over time while subjects receiving placebo will maintain their existing cornstarch needs and their fasting times are unlikely to be modified. Hence, even if TTH at Week 48 is the same as at baseline, if the amount of cornstarch used for CFC at Week 48 is less than that at baseline, it will lead to an increase in TTH/CS at Week 48 compared to baseline. This increase in TTH/CS compared to baseline may be considered as indicative of the endogenous glucose production (ie, by glycogen breakdown due to G6Pase activity) provided by cornstarch. Therefore, an increasing TTH/CS ratio may reflect an increase in G6Pase activity compared to baseline.

Similar supplementary analysis will be performed to assess impact of amount of carbohydrates in the pre-fasting meal on the CFC. MMRM model will be based on change from baseline in TTH/Carbs where TTH/ Carbs denotes the gain in TTH (in minutes) per gram of carbohydrates (including carbohydrates from meal and cornstarch) used prior to CFC.

6.5.3.2.2. Change from Baseline to Week 48 in Percentage of Glucose Values in the range of 70-120 mg/dL

The following non-inferiority hypothesis will be tested for this endpoint with non-inferiority margin of 15%:

H0: Mean Change from Baseline to Week 48 in percentage of glucose values in the range of 70 to 120 mg/dL [3.9 to 6.7 mmol/L] in DTX401 Group – mean change from Baseline to Week 48 in percentage of glucose values in the range of 70 to 120 mg/dL [3.9 to 6.7 mmol/L] in Placebo Group ≤ -15%

H1: Mean Change from Baseline to Week 48 in percentage of glucose values in the range of 70 to 120 mg/dL [3.9 to 6.7 mmol/L] in DTX401 Group – mean change from Baseline to Week 48 in percentage of glucose values in the range of 70 to 120 mg/dL [3.9 to 6.7 mmol/L] in Placebo Group > -15%

The derivation of NI margin of 15% for testing this endpoint is included in Appendix 5.

MMRM Model will include change from baseline in percentage of glucose values in the range of 70-120 mg/dL at Weeks 24, 30, 36, 42, and 48. For each visit, the percentage of glucose values in the range of 70-120 mg/dL will be calculated and the change from baseline will be included in the analysis. Percentage of values is calculated as the number of glucose values in the range of 70-120 mg/dL divided by the total number of glucose measurements recorded. Baseline percentage of glucose values in the range of 70-120 mg/dL is calculated as the percentage of glucose values in the range of 70-120 mg/dL prior to the Baseline (Day 0) visit.

The model will include the treatment group, and Baseline percentage of glucose values in the range of 70-120 mg/dL, stratification factors as captured in IRT, study visits, and interaction of visit by treatment as covariates.

The estimation of the Week 48 contrast (DTX401 – Placebo) will provide the Least Squares Mean estimate of the difference between DTX401 and Placebo. The 97.5% 1-sided confidence interval for the difference in mean change from baseline in percentage of glucose values in 70-120 mg/dL range between DTX401 and Placebo will be calculated. Statistical non-inferiority will be demonstrated if a lower bound of that confidence interval is greater than -15%.

Only if non-inferiority is established, the following hypothesis will be tested at 0.05 alpha to test for superiority. The 2-sided p-value for the comparison and the 95% confidence interval for the estimated difference between DTX401 and Placebo will be provided based on the MMRM model described above.

H0: Mean change from Baseline to Week 48 in percentage of glucose values in the range of 70 to 120 mg/dL [3.9 to 6.7 mmol/L]) in DTX401 Group is same as mean change from Baseline to Week 48 in percentage of glucose values in the range of70 to 120 mg/dL [3.9 to 6.7 mmol/L]) in Placebo Group

H1: Mean change from Baseline to Week 48 in percentage of glucose values in the range of 70 to 120 mg/dL [3.9 to 6.7 mmol/L]) in DTX401 Group is different than mean change from Baseline to Week 48 in percentage of glucose values in the range of 70 to 120 mg/dL [3.9 to 6.7 mmol/L]) in Placebo Group

Additionally, the percentage of glucose values in the range of 70-120 mg/dL and the change from baseline at all scheduled time points will be summarized by the treatment group. Plots of mean percentage of glucose values in 70-120 mg/dL by study visits will be presented by treatment group.

Both continuous glucose monitoring (CGM) data and capillary glucose measurements (SMBG) will be independently used to analyze this endpoint. Two analyses will be performed:

(1) Using CGM data collected in 4 weeks prior to the visit and

(2) Using SMBG data collected in 1 week prior to the visit

The main analysis (i.e., MMRM model and hypothesis testing) for this endpoint will be based on CGM data (1), while analysis based on SMBG measurements (2) will be considered as supplementary analysis and will be used to independently validate conclusions of main analysis.

Additional supplementary analysis will include MMRM analysis based on CGM data collected in 4 weeks prior to the visit but including CGM data from only those weeks where adequate corroboration between CGM and SMBG values (as described in Appendix 1) is observed.

All supplementary analyses will include nominal p-values.

6.5.3.3. Sensitivity Analyses for Secondary Efficacy Endpoints

Missing data:

The strategy to handle the missing endpoint data in the main analysis for key secondary efficacy endpoints is based on ITT principle (treatment policy strategy) i.e., no imputation of missing data

will be performed. Sensitivity analyses for missing data will include imputing missing data for visits included in MMRM with MI technique which imputes data under MAR assumption. SAS PROC MI will be used to create multiply imputed datasets (e.g., 100 datasets) using Markov-Chain Monte Carlo (MCMC). MMRM will be performed on each complete dataset, and results (p-values) will be combined using PROC MIANALYZE.

ICEs:

ICEs for key secondary efficacy endpoints are described in Table 3. The strategy to handle the ICEs in the main analysis of key secondary efficacy endpoints will be based on ITT principle (i.e., treatment policy strategy) i.e., no adjustment will be made to the endpoint values impacted by ICEs and MMRM will be performed with data as-is.

Sensitivity analysis will include hypothetical strategy to handle ICEs i.e., setting data impacted by ICEs to missing and imputing missing values using Multiple Imputation method. SAS PROC MI will be used to create multiply imputed datasets (e.g., 100 datasets) using Markov-Chain Monte Carlo (MCMC) and complete datasets will be analyzed with MMRM model. Results (p-values) from the analyses will be combined using PROC MIANALYZE.

Change in cornstarch data collection from 4 weeks to 1 week prior to the visit during the study: As per schedule of events included in the original Protocol, Protocol Amendment 1 and Protocol Amendment 2, subjects were required to complete cornstarch entries daily for 4 weeks prior to the scheduled visit. That requirement was changed later during the study in the Protocol Amendment 3 to ease data entry burden on subjects and it required subjects to complete cornstarch entries daily for only 1 week prior to the visit, except for Baseline and Week 48 visit where requirement of data collection for 4 weeks prior to the visit was not changed. Hence, it is likely that for a subject, cornstarch intake may have been collected over 4 weeks at some earlier visits in the study and for only 1 week at some later visits in the study.

To maintain consistency in the use of cornstarch data across all subjects in the analysis, only data collected during 1 week prior to the visit will be used (except for Baseline and Week 48 where 4 weeks of data will be used) in the main analysis of key secondary efficacy endpoint of change from baseline in total number of daily doses of cornstarch i.e., average of total daily cornstarch doses over 7 days prior to the visit will be used (except for Baseline and Week 48 where average over 28 days will be used) as the number of total cornstarch doses at a visit. As a sensitivity analysis, all available data in 4 weeks prior to the visit will be used i.e., average of total daily cornstarch doses available over 28 days prior to the visit will be used as the number of total daily cornstarch doses at a visit. This analysis will assess robustness of results of main analysis to changes in cornstarch data collection strategy during the study.

Change in CFC schedule during the study: As per schedule of events included in the original Protocol, and Protocol Amendment 1, CFC was scheduled at Week 12, 24, 36, and 48 during the PEAP. In the Protocol Amendment 2, CFC at Week 12 was removed and in the Protocol Amendment 3, CFC at Week 24 was removed to ease CFC burden on subjects. Hence, it is likely that for some subjects who were enrolled earlier in the study, CFC data at Week 12 and 24 is available and for some subjects who were enrolled later in study that data is not available.

To maintain consistency in the use of CFC data across all subjects in the analysis, only CFC results at Week 36 and Week 48 will be used in the MMRM model in the main analysis of secondary efficacy endpoints based on CFC. As a sensitivity analysis, available data at Week 24

will also be used in the MMRM model in addition to data at Week 36 and Week 48. This analysis will assess robustness of results of main analysis to changes in CFC schedule during the study.

Data at Week 12 will not be used in the main or sensitivity analysis as that was never planned to be included in the MMRM model due to the anticipated steroid effect at that visit.

6.5.3.4. Subgroup analyses

Subgroup analyses for all secondary endpoints will be performed for subgroups described in Section 3.5.

Descriptive statistics of secondary endpoints will be provided within each subgroup. Further, an estimate of the treatment difference based on the MMRM model with corresponding 95% CI will be presented within each subgroup for exploratory purposes.

6.5.3.5. Supplementary analyses

Analyses for all secondary endpoints will be performed for PP analysis set to evaluate the impact of important protocol deviations on secondary efficacy results. Nominal 2-sided p-value will be provided.

Correlation and linear regression analysis to explore the relationship between percent reduction in cornstarch intake from baseline to Week 48 and PGIC at Week 48 will be performed. Plot of mean percent reduction in daily cornstarch intake from baseline to Week 48 by PGIC at Week 48 will be plotted by treatment group. Plot of mean change from baseline in time to hypoglycemia during the CFC at Week 48 by PGIC at Week 48 will be plotted by treatment group.

6.5.4. Analysis of Other Efficacy Endpoint(s)

For each efficacy endpoint included below, CMH test will be performed to compare the endpoints at Week 48, and the estimate of difference in the percentages and 95% confidence interval corresponding to the comparison of interest will be provided for exploratory purpose. Additionally, counts and percentages by treatment group at each scheduled visit will be provided for each endpoint.

- Percentage of subjects with > 75% reduction from baseline in total daily cornstarch.
- Percentage of subjects who do not require cornstarch during the day. If average number of daytime doses (rounded off to nearest integer) in 1 week prior to the visit (or 4 weeks for Baseline and Week 48 visit) is 0, the subject is considered as not requiring cornstarch during the day.
- Percentage of subjects who do not require overnight cornstarch. If average number of nighttime doses (rounded off to nearest integer) in 1 week prior to the visit (or 4 weeks for Baseline and Week 48 visit) is 0, the subject is considered as not requiring overnight cornstarch.

For each efficacy endpoint included below, the MMRM method will be used to analyze the change from baseline to Week 48. Data from Week 24 through Week 48 will be included in the MMRM model. If data is provided only at Week 48 in this period for any endpoint (e.g., GSD FAD data), a linear model will be used which will include fixed effect of treatment group, and

stratification factors and baseline value as covariates. The contrast between groups at Week 48 will provide the least squares mean and 95% confidence interval corresponding to the comparison of interest, for exploratory purpose. Additionally, a summary of raw values and changes from baseline by treatment group at each scheduled visit will be provided for each endpoint.

- Total number of hypoglycemic events (<54 mg/dL) in the 4 weeks prior to the visit based on CGM data. A hypoglycemic event is defined as at least 2 sensor values < 54 mg/dL that are at least 15 minutes apart with no intervening values >54 mg/dL. The end of the hypoglycemic event is defined as 2 sensor values > 70 mg/dL that are at least 15 minutes apart with no intervening values < 70 mg/dL.
- Total duration of hypoglycemic events in the 4 weeks prior to the visit based on CGM data where hypoglycemic events is defined as above.
- Time spent below 70 mg/dL (3.9 mmol/L) in the 4 weeks prior to the visit based on CGM data.
- Plasma Uric Acid
- CFC related lab parameters:

  – Pre-CFC whole blood lactate and rate of rise of lactate during the CFC. Pre-CFC analyte value is the value collected nearest to the start time of the CFC and within +/- 30 minutes of start time of CFC. End-CFC analyte value is the value collected nearest to the end time of the CFC and within +/- 30 minutes of end time of CFC. Peak analyte value during the CFC is the maximum analyte value observed any time from the start time of CFC to the end time of the CFC and including pre-CFC and end-CFC values. Rate of rise of lactate will be calculated as the slope between the pre-CFC lactate value and peak lactate value during the CFC, expressed as mmol/L/h.

  – Pre-CFC plasma glucose and rate of decline of glucose during the CFC. Trough analyte value during the CFC is the minimum analyte value observed any time from the start time of CFC to the end time of the CFC, including pre-CFC and end-CFC values (defined above). Rate of decline of glucose will be calculated as the slope between the pre-CFC glucose value and trough glucose value during the CFC, expressed as mg/dL/h.

  – Pre-CFC plasma alanine

  – Peak values of glucagon, cortisol, growth hormone, IGFBP1, and ketones during the CFC

  – Trough values of insulin and C-peptide during CFC

  – Difference between pre-CFC and end-of-CFC values for glucagon, cortisol, growth hormone inhabiting hormone, IGFBP1, ACTH, fatty acids, ketones, insulin, and C-peptide.

  – Difference between the end-of-CFC and post-CFC parameters may also be explored for selected lab parameters. Post-CFC value is the earliest value

collected after 30 minutes post end of CFC and with same date as the end date of CFC.

- Percent of subjects with inappropriately high insulin level or inappropriately low cortisol levels during the CFC will be presented. Insulin response is deemed inappropriate if ≥ 25 µIU/L at the start of the challenge or > 2µIU/L at the end of the challenge. Cortisol response is deemed inappropriate if < 10 µg/dL at the end of the challenge.

- Time between a glucose level of 70 mg/dL (3.9 mmol/L) to < 54 mg/dL (3.0 mmol/L) during the CFC from baseline. This will be calculated as difference between latest time during CFC when glucose value is at least 70 mg/dL and earliest time during CFC when glucose value is less than 54 mg/dL. This will be based on venous blood glucose values collected in local lab data.
- Subject's observed overnight fasting time as documented by study investigator in CRF based on assessment of glycemic data and cornstarch information.
- Daily dietary macronutrient intake (fat, protein, carbohydrates, and non-utilizable sugars) and % of total calories from individual macronutrients. Non-utilizable sugars as % of total carbohydrates will also be summarized.
- PGIS assessment scores for Felt Like Had Low Blood Sugar, Health Rating, Severity of GSDIa Symptoms, and Bothered by How GSDIa Impacted Life. Relationship between percent reduction in cornstarch intake at Week 48 and change from baseline in PGIS at Week 48 will be explored. Plot of mean percent reduction in daily cornstarch intake from baseline to Week 48 by change in PGIS at Week 48 will be plotted by treatment group.
- PGIC-CTT assessment score at Week 48. Relationship between percent reduction in cornstarch intake at Week 48 and PGIC-CTT at Week 48 will be explored.
- GSD FAD symptom score(s) and impact score(s). The GSD FAD psychometric validation report will describe score derivation including rules for handling missing data. Relationship between percent reduction in cornstarch intake at Week 48 and change from baseline in GSD FAD scores at Week 48 will be explored.
- Patient-reported health status (summary index score and VAS score) via the EQ-5D-5L/EQ-5D-Y instrument.
- MRU, specifically number of hospitalizations and ER visits. Number of nights hospitalized and number of nights in ICU will be summarized.
- Work and school status and productivity, assessed via questionnaire.

Additional analyses will include following:

- Findings from patient exit interviews will be mainly described in a qualitative manner. If there is any quantitative data available as part of interviews (e.g., subject's guess of treatment received at Day 1 IP administration, rating of importance to reduce amount and/or frequency of cornstarch prior to the start of the trial, overall satisfaction with treatment, treatment preference, treatment convenience, benefit/risk of gene therapy, recommend gene therapy to others, change in day-to-day management of GSDIa since start of the study, change in worry about blood sugar since start of the study), descriptive statistics of these data will be presented. Also, exploratory analyses may be performed to explore the relationship between subject's responses in exit interview and subject's efficacy and safety profile.

6.5.4.1. Supplementary analyses

Supplementary analyses of efficacy endpoints will include following:

- Descriptive analysis of response to individual items in GSD FAD eDiary will be provided.
- Subgroup analyses for GSD FAD score(s) may be performed for subgroups of interest e.g., age groups, BMI groups, race/ethnicity, baseline daily cornstarch intake, Glycosade use (Yes/No) as documented on Prescribed Daily Cornstarch Intake eCRF at baseline visit, on continuous feed (Yes/No) as documented on Prescribed Daily Diet eCRF at baseline, total number of hypoglycemic episodes (<54 mg/dL) in the 4 weeks prior to the baseline visit based on CGM data, sleep disturbance related to GSD1a treatment at baseline as documented on GSDIa specific medical history eCRF, patient reported health status via EQ-5D at baseline.
- The GSD FAD Evening Diary asks subjects if they feel they have had low blood sugar on a given day. SMBG and CGM measurements from that day will be matched with diary responses to determine if subjects were able to detect hypoglycemic episodes that occurred during waking hours. Matching the diary responses and blood glucose values will also help to identify the range of values at which subjects typically experience symptoms. The percentage of subjects who fail to identify episodes or report symptoms on a given day where blood glucose levels indicate hypoglycemia will be reported.
- Before the start of the CFC and after the end of the CFC, the site staff will ask the subject to rate the current severity of their hypoglycemia symptoms (low blood sugar, fatigue/tiredness, weakness, and confusion/brain fog). The presence of symptoms and their severity will be summarized at each scheduled visit. Further, a linear regression model will be used to compare the mean severity of a symptom (e.g., fatigue/tiredness) after the end of CFC at Week 48 between DTX401 and placebo group using severity of that symptom before the start of the CFC at Week 48 as a covariate in the regression model.
- Descriptive analyses of symptomatic hypoglycemia events reported by subject in eDiary in 4 weeks prior to the visit will be provided.

6.6. Safety Analysis

The safety analysis will be performed using the safety analysis set. Safety parameters will include AEs, vital sign measurements, physical examinations, weight, ECGs, liver imaging, safety laboratory assessments, and concomitant medications.

6.6.1. Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and displayed in tables and listings using System Organ Class (SOC) and Preferred Term (PT).

An AE (classified by preferred term) will be considered a treatment emergent adverse event (TEAE) if it occurred on or after the first dose of investigational product and was not present prior to the first dose, or it was present at the first dose but increased in severity during the study.

An AE will be considered a treatment related AE if relationship to investigational product/placebo is marked as "Related" on AE CRF. The Investigator will make a causality assessment about the relationship of each AE to study drug. Serious AEs (SAEs) will be identified as AEs with "Is the event serious?" marked as Yes on AE CRF. The severity of AE will be assessed by Investigator based on Common Terminology Criteria for Adverse Events (CTCAE) version 5.0 (Appendix 3) and captured on AE CRF. If an AE cannot be graded based on CTCAE, the Investigator will assign a severity based on 1 = mild, 2 = moderate, 3 = severe, 4 = life threatening, and 5 = death as per the instructions provided in the study protocol.

Adverse events will be summarized by subject incidence rates, therefore, in any tabulation, a subject will contribute only once to the count for a given AE (SOC or PT). No formal hypothesis-testing analysis of AE incidence rates will be performed.

Subject incidence of AEs prior to the first dose will be tabulated by preferred term. Subject incidence of TEAEs will be tabulated as follows. The number of events may also be presented. Exposure-adjusted incidence rates may also be provided, as appropriate.

- TEAEs by treatment, SOC, and PT (sorted by descending frequency by SOC and PT)
- Treatment related TEAEs by treatment, SOC, and PT (sorted by descending frequency by SOC and PT)
- Serious TEAEs by treatment, SOC, and PT (sorted by descending frequency by SOC and PT)
- Serious treatment related TEAEs by treatment, SOC, and PT (sorted by descending frequency by SOC and PT)
- TEAEs by treatment, PT, and severity (sorted by greatest severity, descending frequency by PT)
- Treatment related TEAEs by treatment, PT, and severity (sorted by greatest severity, descending frequency by PT)
- Grade 3 or 4 TEAEs by treatment, SOC, and PT (sorted by severity, descending frequency by SOC and PT)

- TEAEs leading to deaths by treatment, SOC, and PT (sorted by descending frequency by SOC and PT)
- TEAEs leading to discontinuation of study by treatment, SOC, and PT (sorted by descending frequency by SOC and PT)
- TEAEs leading to discontinuation of study drug by treatment, SOC, and PT (sorted by descending frequency by SOC and PT)
- TEAEs leading to study drug dose interruption by treatment, SOC, and PT (sorted by descending frequency by SOC and PT)
- TEAEs by treatment and PT (sorted by descending frequency of PT)
- Serious TEAEs by treatment and PT (sorted by descending frequency of PT)
- Non-serious TEAEs by treatment and PT (sorted by descending frequency of PT)
- Treatment related TEAEs by treatment and PT (sorted by descending frequency of PT)
- Prednisolone/Placebo related TEAEs by treatment and PT (sorted by descending frequency of PT)
- Study participation related TEAEs by treatment and PT (sorted by descending frequency of PT)
- TEAEs during the CFC by treatment and PT (sorted by descending frequency of PT)

Detailed listings for all AEs, serious TEAEs, Grade 3 or 4 TEAEs, study participation related AEs, AEs leading to the discontinuation of study, and death will also be generated.

6.6.1.1. Important Potential Risks and Adverse Events of Special Interest (AESI)

Vector-induced hepatic effects (e.g., increased aminotransferase levels) following IP administration and infusion related reactions (IRR) including hypersensitivity during or following IP administration have been identified as important potential risks and AESIs associated with DTX401. MedDRA search strategy to identify PTs for these events is included in Appendix 4. IRR will be identified as TEAEs with onset during infusion or up to 6 hours after end of the infusion and with PTs identified by MedDRA search strategy included in Appendix 4.

Subject Incidence of TEAEs of special interest will be tabulated as follows:

- TEAEs of special interest by SOC and PT
- Treatment related TEAEs of special interest by SOC and PT
- Serious treatment related TEAEs of special interest by SOC and PT
- TEAEs of special interest by PT and greatest reported severity in descending frequency
- Treatment related TEAEs of special interest by PT and greatest reported severity in descending frequency

Additionally, relationship between staged infusion (Yes/No) and incidence rate, severity, and seriousness of IRR will be explored.

6.6.1.2. Potential Risk and Adverse Event of Special Interest (AESI)

Adrenal insufficiency is considered a potential risk of study-related procedures associated with use of corticosteroids/placebo and is an AESI. MedDRA search strategy to identify PTs for these events is included in Appendix 4.

Subject Incidence of TEAEs of adrenal insufficiency will be tabulated as follows:

- TEAEs by SOC and PT
- Treatment related TEAEs by SOC and PT
- Serious treatment related TEAEs by SOC and PT
- TEAEs by PT and greatest reported severity in descending frequency
- Treatment related TEAEs by PT and greatest reported severity in descending frequency

6.6.1.3. Potential Risk

Dorsal Root Ganglion and Peripheral Nerve Effect following IP administration has been identified as a potential risk for DTX401. MedDRA search strategy to identify PTs for these events is included in Appendix 4.

Subject Incidence of TEAEs of potential risk will be tabulated as follows:

- TEAEs of potential risk by SOC and PT
- Treatment related TEAEs of potential risk by SOC and PT
- Serious treatment related TEAEs of potential risk by SOC and PT
- TEAEs of potential risk by PT and greatest reported severity in descending frequency
- Treatment related TEAEs of potential risk by PT and greatest reported severity in descending frequency

6.6.1.4. AAV Gene Therapy Class Effects

AAV gene therapy class effects include malignancies (new or worsening of pre-existing malignancies) and thrombotic microangiopathy. The MedDRA search strategies to identify PTs for these events is included in Appendix 4.

Subject Incidence of TEAEs for AAV gene therapy class effects will be tabulated as follows:

- TEAEs by SOC and PT
- Treatment related TEAEs by SOC and PT
- Serious treatment related TEAEs by SOC and PT
- TEAEs by PT and greatest reported severity in descending frequency

- Treatment related TEAEs by PT and greatest reported severity in descending frequency

6.6.2. Clinical Laboratory Parameters

Quantitative laboratory measurements reported as "< X", i.e., below the lower limit of quantification (BLQ) will be converted to X/2 for the purpose of quantitative summaries, but will be presented as recorded, i.e., as "< X" in the listings. Quantitative laboratory measurements reported as "> X", i.e., above the upper limit of quantification (ULQ), will be converted to X for the purpose of quantitative summaries, but will be presented as recorded, i.e., as "> X" in the listings.

In the event of repeat laboratory assessments, the last non missing value will be used. If a sample is analyzed under 2 different panels (Chemistry vs AST/ALT), it is likely that AST and ALT may have same collection time but possibly different results. In such case, values collected under Chemistry panel will be included in the summary tables, but listings will show values collected under both panels.

Descriptive statistics for clinical laboratory values (in SI units) and change from baseline (if applicable) at each assessment time point will be presented by treatment group for the lab parameters included in Table 6.

Samples to assess complement activity and serum tryptase will be taken prior to dosing at Day 1 and Week 48 and only post dosing if the subject experiences an IRR in which case these labs will be measured 30 to 120 minutes after the start of the reaction, and once again at least 24 hours after the symptoms have resolved. Only data collected prior to dosing for these analytes will be summarized and post-dosing data (if any) for events of IRRs will be included in the listing.

Table 6: Clinical laboratory Parameters

Panel	Assessments
Chemistry	ALT, AST, ALP, GGT, LDH, Albumin, total protein, Bilirubin (total, direct, and indirect), BUN, creatinine, Lipid (triglycerides, LDL, HDL, total cholesterol) panel, Sodium, potassium, chloride, carbon dioxide, Glucose, Calcium, Phosphate, Magnesium, Creatine kinase, Uric acid, creatinine clearance/eGFR
Hematology	CBC with differential
Urinalysis	Specific gravity, pH, glucose, protein, Blood and ketones, Microscopic examination (if blood or protein is found)
Coagulation panel	PT/INR, aPTT
24-hour Urine	total protein, microalbumin, creatinine, and Hex4
Complement activity at Baseline	Complement (C3, C4, SC5b-9, CH50, AH50, MBL) and serum tryptase

ALP, alkaline phosphatase; ALT, alanine aminotransferase; aPTT, activated partial thromboplastin time; AST, aspartate aminotransferase; CBC, complete blood count; GGT, gamma-glutamyl transferase; HDL, high density lipoprotein; INR, international normalized ratio; LDH, lactate dehydrogenase; LDL, low density lipoprotein; PT, prothrombin time

Quantitative laboratory measurements will be compared with the relevant laboratory reference ranges in SI units and categorized as:

- Low: Below the lower limit of the laboratory reference range
- Normal: Within the laboratory reference range (upper and lower limit included)
- High: Above the upper limit of the laboratory reference range

The frequency and percentage of subjects who experience abnormal clinical laboratory results (ie, outside of reference ranges) will be presented for each clinical laboratory measurement. A listing of abnormal clinical laboratory measurements will be provided for selected parameters.

Severity of chemistry and hematology laboratory parameters will be graded according to CTCAE version 5.0 (Appendix 3) if severity grading criteria are available for the parameter. Shift tables using CTCAE grades to compare baseline to the worst post-baseline value will be produced for hematology and chemistry laboratory parameters. For some specific parameters with CTCAE grading in both high and low direction (e.g., glucose, potassium, sodium), CTCAE grades in high and low directions will be presented separately, i.e., hyper for higher values of concern and hypo for lower values of concern. Values from post-baseline unscheduled visits will also be considered when the worst post-baseline value is derived.

6.6.3. Vital Signs

Descriptive statistics for vital signs (heart rate, seated systolic and diastolic blood pressures, and respiratory rate) and changes from baseline values at each assessment time point and at the end of study will be presented.

Vital sign values will be considered clinically significant if they meet both the observed-value criteria and the change-from-baseline criteria listed in Table 7. The number and percentage of subjects with clinically significant post-baseline values will be tabulated. The percentages will be calculated relative to the number of subjects with available baseline values and at least 1 post-baseline assessment. The numerator will be the total number of subjects with available baseline values and at least 1 clinically significant post-baseline value. A listing of clinically significant vital signs will be provided.

Table 7: Criteria for Clinically Significant Vital Signs

Parameter	Flag	Criteria			
		Observed Value			Change from Baseline
		Adults	Children 8-12 years	Children 13-17 years	All
Systolic blood pressure, mm Hg	High	≥ 180	≥ 135	≥ 150	Increase of ≥ 20
	Low	≤ 90	≤ 80	≤ 90	Decrease of ≥ 20
Diastolic blood pressure, mm Hg	High	≥ 105	≥ 90	≥ 100	Increase of ≥ 15
	Low	≤ 50	≤ 40	≤ 50	Decrease of ≥ 15
Heart rate, bpm	High	≥ 120	≥ 130	≥ 120	Increase of ≥ 15
	Low	≤ 60	≤ 60	≤ 60	Decrease of ≥ 15

Height, weight, BMI, and changes from baseline (not applicable for height) will be summarized for all adult subjects. Percentiles and Z-scores for height, weight, and BMI, and changes from baseline will be summarized for all pediatric subjects. Percentiles and Z-scores will be calculated using the SAS code provided by the Centers for Disease Control/National Center for Health Statistics (CDC/NCHS) which calculates CDC percentiles and z-scores based on the 2000 CDC growth charts (http://www.cdc.gov/growthcharts/cdc_charts.htm). The reference population is children examined in NCHS studies from 1963-65 to 1988-94.

Physical examination results (determined as normal, abnormal and not clinically significant, or abnormal and clinically significant by the investigator) at each scheduled visit will be summarized. Tanner pubertal stage development data will be summarized.

6.6.4. Electrocardiogram

ECG results will be summarized descriptively, including the number and percent of subjects with normal, non-clinically significant abnormal, and clinically significant abnormal results as assessed by the study investigator at Baseline and each study visit.

6.6.5. Liver Imaging

Liver imaging (Ultrasound, MRI, CT scan) results will be summarized descriptively, including the number and percent of subjects with normal and abnormal results as assessed by the study investigator at Baseline and each study visit. Abnormal findings (hepatomegaly and hepatic adenomas) will be listed (if any) and may be summarized. Further details related to hepatic adenomas will be listed including number of adenomas, maximum diameter of adenomas and growth rate.

6.6.6. Other Safety Parameters

Measurement of AAV8 binding antibody immunoglobulin G (IgG), measurement of anti-G6Pase antibodies, and Cell-mediated immune response to G6Pase (ELISpot) will be summarized descriptively by visit and treatment group and/or provided in the listings.

G6PC genotyping results at screening will be listed. Virus screening results, serum pregnancy test results, urine pregnancy test results, and 9-panel urine drug panel test results will be listed.

6.6.7. COVID-19 Impact on Subject Participation

The study CRFs will capture COVID-19 related AEs and reasons for study discontinuation, missed/delayed visits, and missed assessments due to COVID-19 as “COVID-19 related reason”. Study data listings will include this information e.g., listing of AEs will include any AEs related to COVID-19 and listing of protocol deviations will include deviations related to missed visits and missed assessments due to COVID-19. These listings will be used to assess any potential impact of COVID-19 on subjects’ study participation.

7. CHANGES TO ANALYSIS SPECIFIED IN PROTOCOL

There are no changes to analyses specified in the protocol.

8. LIST OF PLANNED TABLES AND FIGURES

A list of planned tables, figures, and listings will be included in the shell document which will be maintained separately.

9. LITERATURE REFERENCES

DHHS (2017) *Common terminology criteria for adverse events (CTCAE) Version 5.0 (27 November 2017)*. Available at: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm.

Griffiths P, Williams A and Brohan E (2022) How do the number of missing daily diary days impact the psychometric properties and meaningful change thresholds arising from a weekly average summary score? *Qual Life Res* 31(12): 3433–3445.

Mallinckrodt CH, Lane, P.W., Schnell, D. et al. (2008) Recommendations for the Primary Analysis of Continuous Endpoints in Longitudinal Clinical Trials. *Drug Information Journal* 42(4): 303-319.

Rossi A, Rutten MGS, van Dijk TH, et al. (2022) Dynamic Methods for Childhood Hypoglycemia Phenotyping: A Narrative Review. *Front Endocrinol (Lausanne)* 13: 858832.

Shah VN, Laffel LM, Wadwa RP, et al. (2018) Performance of a Factory-Calibrated Real-Time Continuous Glucose Monitoring System Utilizing an Automated Sensor Applicator. *Diabetes Technol Ther* 20(6): 428-433.

10. APPENDICES

APPENDIX 1. HANDLING OF MISSING AND INCOMPLETE DATA

Missing clinical outcome data can occur for multiple reasons, including missed subject visits and scales or measures with missing item scores. Missing and incomplete data will be identified through the data quality review plan for this study. Missing and incomplete data will be identified for investigation, and possible resolution, by Data Management prior to the study database lock or snapshot.

Unless specified otherwise, only the observed data (not imputed data) will be presented in listings.

Missing Medical History Related Dates

To impute a missing medical history related date, the following rules will be applied:

- If only day is missing, impute 1.
- If month is missing, impute January 1st.
- If year is missing, then no imputation will be done; the date will be missing.

If the imputed date is earlier than the birth date, then the birth date will be used.

Missing Birth Dates

To impute missing birth date, the following rules will be applied:

- If day is missing, impute 15.
- If month is missing, impute June.
- If year is missing, then no imputation will be done; the date will be missing.

If the imputed date is later than any study visit date/observed adverse event start date/ observed concomitant medication start date, then earliest available visit date/adverse event start date/ concomitant medication start date will be used without changing observed information.

Missing Date Imputation for Adverse Events and Concomitant Medications

The following conventions will be used to impute missing portions of dates for adverse events and concomitant medications. Note that the imputed values outlined here may not always provide the most conservative date.

Missing Start Dates

- If the day is unknown, then:
  - If the month and year match the first dose of investigational product start date month and year in this study, then impute the day of the first dose date.
  - Otherwise, assign the first day of the month.
- If the month is unknown, then:
  - If the year matches the year of the first dose of investigational product date in this study, then impute the month and day of the first dose date in this study.

  - Otherwise, assign ‘January’

- If the year is unknown, then the date will not be imputed and will be assigned a missing value, unless the end date (observed or imputed as described below) is on or after the first IP dose date or the event is ongoing in which case the start date will be imputed as the first IP dose date.

If the imputed date is earlier than birth date, then birth date will be used. If the imputed start date is later than the end date, then the start date will be set as the same date as the end date.

Missing Stop Dates and not ongoing (End dates will be imputed only if complete end date is required for the analysis)

- If the day is unknown, then assign the last day of the month.
- If the month is unknown, then assign ‘December.’
- If the year is unknown, then the date will not be imputed and will be assigned a missing value, and the event will be considered ongoing. If the AE has been recorded as resolved/recovered, all efforts should be made to obtain the date from the Investigator.

If the resulting end date is after the date of study completion / discontinuation/ data cutoff, set the imputed end date as close to the date of study completion / discontinuation/ data cutoff as possible without overwritten existing information.

Missing Causal Relationship to Investigational Product for Adverse Events

If the causal relationship to the investigational product is missing for an AE that started on or after the date of the first dose of double-blind investigational product, a causality of “related” will be assigned. The imputed values for causal relationship to investigational product will be used for the incidence summary; the values will be shown as missing in the data listings.

Missing Data for Efficacy Endpoints

Only observed efficacy data will be used in the main analysis i.e., no imputation of missing efficacy data will be performed in the main analysis. Missing efficacy endpoint data are addressed by using the MMRM, which assumes data missing at random (MAR) i.e., missing data can be modelled using available, observed values and that it does not depend on the values of unobserved data.

Sensitivity analyses will be performed for primary efficacy endpoint and key secondary efficacy endpoints by handling missing data differently than handled in the main analysis to provide a robust assessment of the results of main analysis.

Handling of Spurious CGM Data

CGM values less than 40 mg/dL or greater than 400 mg/dL will not be included in the analysis because they are not considered as reliable values being outside the Dexcom G6 reportable range of 40-400 mg/dL.

If the CGM device is malfunctioning or not worn properly by the subject, it may lead to spurious CGM measurements. CGM measurements from weeks where adequate corroboration between

the SMBG values and CGM values is not observed will be excluded from the sensitivity analysis. The main analysis will include all CGM data irrespective of the corroboration between the SMBG values and CGM values.

Adequate corroboration is defined as the median absolute relative difference (MRD) as <20% and mean bias between –15 mg/dL and 15 mg/dL among the pairs of SMBG and CGM data collected around similar time in that week, provided that at least 20 such pairs are available in that week. These thresholds are based on reported bias and mean absolute relative difference (MARD) in studies comparing CGM vs SMBG values in adult GSDIa patients (Rossi et al., 2022) and CGM vs blood glucose values determined by the laboratory comparator analyzer in diabetic patients (Shah et al., 2018). MRD is calculated as the median of absolute relative difference (ARD) values where ARD is calculated as the absolute difference in paired value of SMBG and CGM divided by the SMBG value. Bias is the difference between paired value of SMBG and CGM.

Handling of Duplicate Collection Time in CGM Data

If there are multiple CGM values collected at same local time (due to device malfunctioning or time adjustments due to daylight savings), only the record with latest transcription time will be included in the analysis.

Handling of Duplicate Collection Time in Cornstarch Data in eDiary

There can be instances where subject may have entered different cornstarch doses but at the same local date/time in eDiary due to data entry error. In these cases, only the record with latest transcription time will be included in the analysis.

Handling of Duplicate Visit Names and Assessment Dates in eDiary GSD FAD Data

FAD data will primarily be entered into the eDiary by subjects for 7 consecutive days prior to the scheduled visit (except at the Baseline visit when the GSD FAD is completed for 14 consecutive days ahead of the visit) and the source of this data will be documented as “device” in the clinical database. In certain circumstances, subjects may complete the GSD FAD on paper. Site coordinators will manually enter the paper-collected data into the eDiary via the portal. In these cases, the source of data will be documented as “paper” in the clinical database.

Due to challenges with eDiary visit scheduling and the need to collect some of the GSD FAD data on paper, there are instances of duplicate data. Rules for the handling of these data in the analysis include:

- In the event of a duplicate entry from the device and paper with the same visit name (e.g., Week 48 Day Minus 3) but different date, the device data will be used.
- In the event of a duplicate entry with the same visit name (e.g., Week 48 Day Minus 3) and different dates coming from the same source (paper or device), the data entered for the later date will be used.

APPENDIX 2. VISIT WINDOWS

Analysis windows described below will be used for mapping scheduled or unscheduled visits to analysis visit.

After mapping, if there is more than one visit in the same analysis window, the scheduled visit will be used if available; if there is no scheduled visit in the same window, the mapped visit closer to the target assessment day will be used. If more than one visit has equal distance to the target day, then the later one will be used. If more than one visit is on the same day, the time or the sequence number will be used to select the later record. For listings and shift tables, all data points will be included.

Analysis windows will be used for all efficacy assessments and selected safety assessments (central lab chemistry, hematology, urinalysis, and coagulation). If the visit is not included in the table below (e.g., Week 8 through Week 11), no analysis window is defined for that visit and all data collected for that visit irrespective of the timing of the visit will be included in the analysis.

For identifying cornstarch and diet assessments collected in eDiary corresponding to a visit, if the scheduled visit is in the analysis window for the visit, the corresponding visit date will be used as an anchor date and data reported by the subject in 1-4 weeks (as applicable) prior to the anchor date will be included in the analysis for that visit. If a scheduled visit is not available, visit date for an unscheduled visit available in the analysis window which is nearest to target day will be used as an anchor date. If no scheduled visit or an unscheduled visit is not available in the analysis window, target day for the visit will be used as an anchor date. Same method will be used for identifying CGM and SMBG data corresponding to a visit.

For FAD, PGIS, and PGIC assessments collected in eDiary, provided that the scheduled visit is in the analysis window, all data reported by the subject corresponding to that visit will be included in the analysis.

Analysis Visit	Target Day[a]	Protocol-Specified Window	Analysis Window
Baseline	0	0	Study Day<=Day 1 and prior to dosing
Week 12	84	± 3	[Target day-20, Target Day +21]
Week 24	168	± 14	[Target day-42, Target Day +21]
Week 30	210	± 14	[Target day-20, Target Day +21]
Week 36	252	± 14	[Target day-20, Target Day +21]
Week 42	294	± 14	[Target day-20, Target Day +21]
Week 48	336	+ 7	For subjects who received second IP at Week 48 visit [Target Day-20, Week 48 Dosing Visit Date-1] For subjects who did not receive second IP at Week 48 visit [Target Day-20, Target Day +42] For subjects who discontinued after PEAP, Study Day ≥ Target Day-20
Week 60	420 (IP2+83)	± 3	[Target day-20, Target Day +21]
Week 72	504 (IP2+167)	± 14	[Target day-42, Target Day +21]

Analysis Visit	Target Day[a]	Protocol-Specified Window	Analysis Window
Week 78	546 (IP2+209)	± 14	[Target day-20, Target Day +21]
Week 84	588 (IP2+251)	± 14	[Target day-20, Target Day +21]
Week 90	630 (IP2+293)	± 14	[Target day-20, Target Day +21]
Week 96	672 (IP2+335)	+ 7	[Target day-20, Target Day +42]
Week 108	756 (IP2+419)	± 14	[Target day-41, Target Day +42]
Week 120	840 (IP2+503)	± 14	[Target day-41, Target Day +42]
Week 132	924 (IP2+587)	± 14	[Target day-41, Target Day +42]
Week 144	1008 (IP2+671)	± 14	[Target day-41, Target Day +42]

IP2= Second IP date at Week 48

For visits up to Week 48, study day included in the table will be used as target day for calculating analysis window.

For visits after Week 48, for DTX401 group or if subject does not get second IP at Week 48, study day included outside the parenthesis will be used as target day for calculating analysis window. For Placebo group subject who gets second IP at Week 48, target day will be calculated using second IP date (IP2) as included in the parenthesis.

APPENDIX 3. CTCAE TOXICITY GRADES

Refer to CTCAE Version 5.0 (DHHS, 2017) at
https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm.

APPENDIX 4. MEDDRA SEARCH STRATEGIES FOR IMPORTANT POTENTIAL RISKS, ADVERSE EVENTS OF SPECIAL INTEREST (AESI), POTENTIAL RISKS, AND AAV GENE THERAPY CLASS EFFECTS

Following MedDRA search strategy will be used to identify preferred terms for the important potential risk and AESI of Vector-induced hepatic effects:

- Drug related hepatic disorders- comprehensive search SMQs (with all 4 sub-SMQs: Cholestasis and jaundice of hepatic origin; Drug related hepatic disorders – severe events only (includes 4 sub-SMQs); Liver related investigations, signs and symptoms; Liver related coagulation and bleeding disturbances).

Following MedDRA search strategy will be used to identify preferred terms for the important potential risk and AESI of Infusion-Related Reactions (IRR):

- Hypersensitivity SMQ (broad + narrow terms) and Anaphylactic reaction SMQ (broad + narrow terms)

Following MedDRA search strategy will be used to identify preferred terms for the potential risk and AESI of Adrenal insufficiency:

- Adrenal insufficiency, Adrenocortical insufficiency acute, Secondary adrenocortical insufficiency, Adrenal suppression, Acute adrenal crisis, Steroid withdrawal syndrome, Glucocorticoid deficiency, Cortisol decreased, Cortisol abnormal, Blood corticotrophin decreased, Blood corticotrophin abnormal, Adrenocorticotropic hormone deficiency , ACTH stimulation test abnormal

Following MedDRA search strategy will be used to identify preferred terms for the potential risk of Dorsal Root Ganglion (DRG) and Peripheral Nerve Effects:

- HLGT Peripheral Neuropathies

Following MedDRA search strategy will be used to identify preferred teams for the AAV gene therapy class effect of malignancies (new or worsening of pre-existing malignancies):

- Malignancies SMQ [(broad + narrow terms) with sub-SMQs of Malignancy related conditions, Malignancy related therapeutic and diagnostic procedures, Malignant or unspecified tumors, Tumor markers]

Following MedDRA search strategy will be used to identify preferred teams for the AAV gene therapy class effect of thrombotic microangiopathy:

- Embolic and Thrombotic Events SMQ [(broad + narrow terms) with sub-SMQs of Embolic and thrombotic events, arterial; Embolic and thrombotic events, venous; Embolic and thrombotic events, vessel type unspecified and mixed arterial and venous] and
- Hemolytic disorder SMQ (broad + narrow terms)

APPENDIX 5. DERIVATION OF NON-INFERIORITY (NI) MARGINS

Derivation of NI margin for the endpoint of "change from baseline in percent of glucose values in <70 mg/dl range"

Non-inferiority (NI) margin for testing this endpoint is set up referencing FDA guideline ("Non-Inferiority Clinical Trials to Establish Effectiveness", dated November 2016) and EMA guideline ("Guideline on the choice of non-inferiority margin, EMEA/CPMP/EWP/2158/99"). To derive NI margin for an endpoint, guidelines recommend following steps:

1. Estimate the minimal effect of the active control relative to placebo (say M1), based on past randomized controlled trials; specifically lower 95% confidence limit of the treatment effect (i.e., active control – placebo).
2. Set non-inferiority margin (M2) as some clinically relevant portion of M1 which represents clinical judgment about the amount of the active control effect that must be retained in the test drug. Usually, 50% of M1 is set as NI margin.

For this study, M1 will need to be estimated based on 95% confidence interval of difference in effects between active comparator (i.e., GSDIa patients using cornstarch) and placebo (i.e., GSDIa patients using no cornstarch). Without cornstarch, GSDIa patients would be expected to die of severe hypoglycaemia and lactic acidosis, hence % of glucose values in <70 mg/dL for such patients will be very high, hypothetically it will be close to 100%. Historical data on glycaemic control in GSDIa patients who manage the disease with cornstarch indicates that % of glucose values in <70 mg/dL is usually < 10% (see results described below). Hence, M1 based on the difference between GSDIa patients using no cornstarch versus GSDIa patients using cornstarch will be very large i.e., close to 90% and M2 as 50% of M1 will also be large. Hence, clinical judegment was sought from multiple KOLs to determine the clinically relevant margin. The KOLs determined that 10% difference in % of glucose values in <70 mg/dL would be clinically reasonable and within the threshold for clinical stability. Based on this, 10% margin was chosen for testing this endpoint.

This margin appears suitable for this study based on the precision of difference in effects between DTX401 group and Placebo group (GSDIa patients using cornstarch) expected in this study as per the simulations described below.

For simulations, assumptions are based on the historical data on glycaemic control measured by CGM in GSDIa patients who manage the disease with cornstarch. Following are the only studies available which report CGM data in GSDIa patients.

- Observational study of 15 GSDIa patients of age 8-48 years (DTX401-CL001),
- Baseline (pre-dosing) data of 4 adult GSDIa patients of age 22-42 years in an open-label phase 1/2 study of DTX401 (401GSDIA01), and
- Prospective study of CGM in 10 adult GSDIa patients of age 17-48 years (Rossi et al., 2022)

The mean and standard deviation (SD) of % glucose values in <70 mg/dL observed in these studies are included below.

% glucose values in <70 mg/dL	DTX401-CL001 (N=15)	401GSDIA01 (N=4)	(Rossi et al., 2022) (N=10)
Mean	3.1	8.6	3.5
SD	2.6	12.6	3.6

Based on pooled analysis of data from these 3 studies, the mean (SD) of % of glucose values <70 mg/dL is estimated as 4% (6%).

Using these estimates 1000 datasets were simulated, each containing % of glucose values in <70 mg/dL at baseline and Week 48 for 25 subjects in DTX401 group and 25 subjects in Placebo group (Group using only cornstarch). For both DTX401 and Placebo group, baseline and Week 48 values were simulated from bivariate normal distribution with mean at baseline as 4%, mean at Week 48 as 4%, SD at baseline as 6%, SD at Week 48 as 6%, and correlation between baseline and Week 48 as 0.4. This is under the assumption that there is no difference between 2 groups at baseline and both groups maintained the same level of glycemic control as at baseline over time. The estimate of the difference in mean reduction from baseline to Week 48 (reduction is defined as Week 48 value - baseline value) in the % of glucose values in <70 mg/dL and the corresponding 95% CI is calculated for each dataset using 2-sample t-test. Figure 2 shows the histogram of the difference in mean reductions of 2 groups based on estimates in 1000 datasets. It shows that the lower limit for difference can be up to –6%. Figure 3 shows the histogram of the lower 95% confidence limit for the difference in the mean reductions of 2 groups based on estimates in 1000 datasets. It shows that the lower limit of 95% CI for the difference can be up to –10%.

This indicates that even if no difference is assumed between 2 groups, because of the inherent variability in the data, the lower limit of the confidence interval of the difference between treatment groups can be up to -10%. Based on this, 10% is chosen as a margin for testing non-inferiority (NI) of DTX401 to Placebo for this endpoint under the assumption of SD as 6%. If the study analysis results show much larger SD, showing statistically compelling evidence of non-inferiority may not be feasible. In that case demonstration that the reduction in cornstarch intake in the DTX401 group is not accompanied by any deterioration in glycemic control will be based not only on this non-inferiority test, but on the totality of information generated in respect of changes in glucose.

Figure 2: Histogram of Difference in Mean Reduction in % Glucose Values <70 mg/dL Between DTX401 and Placebo

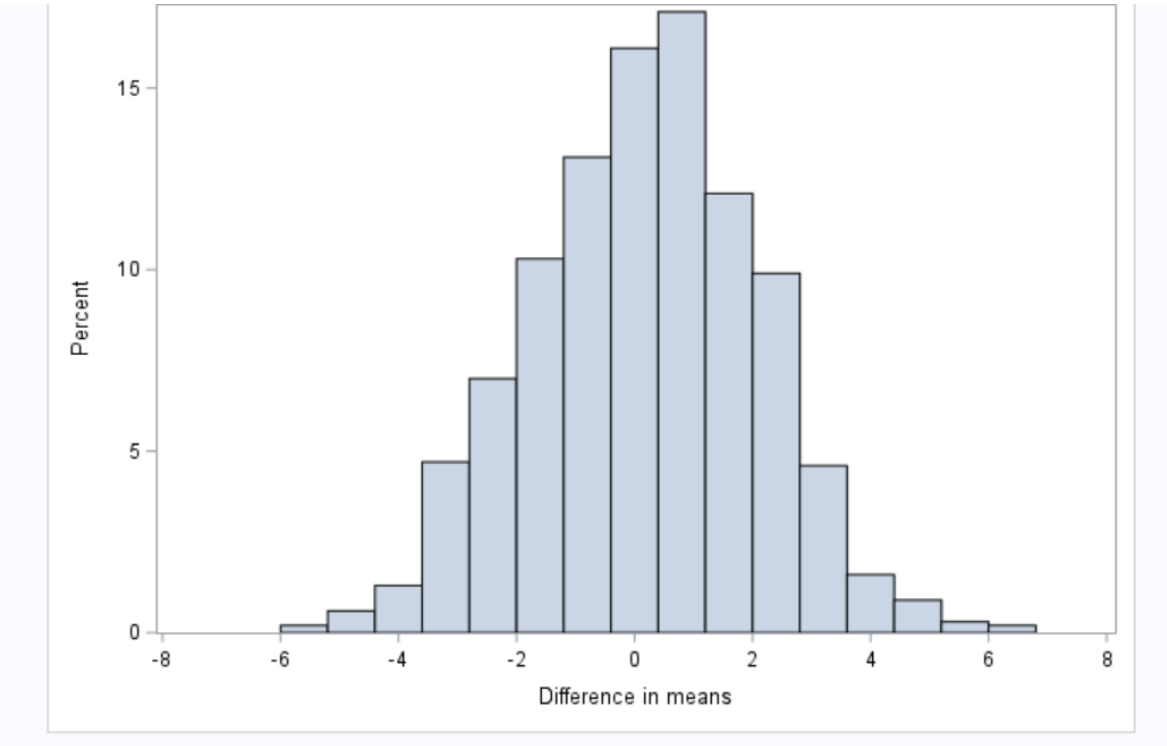
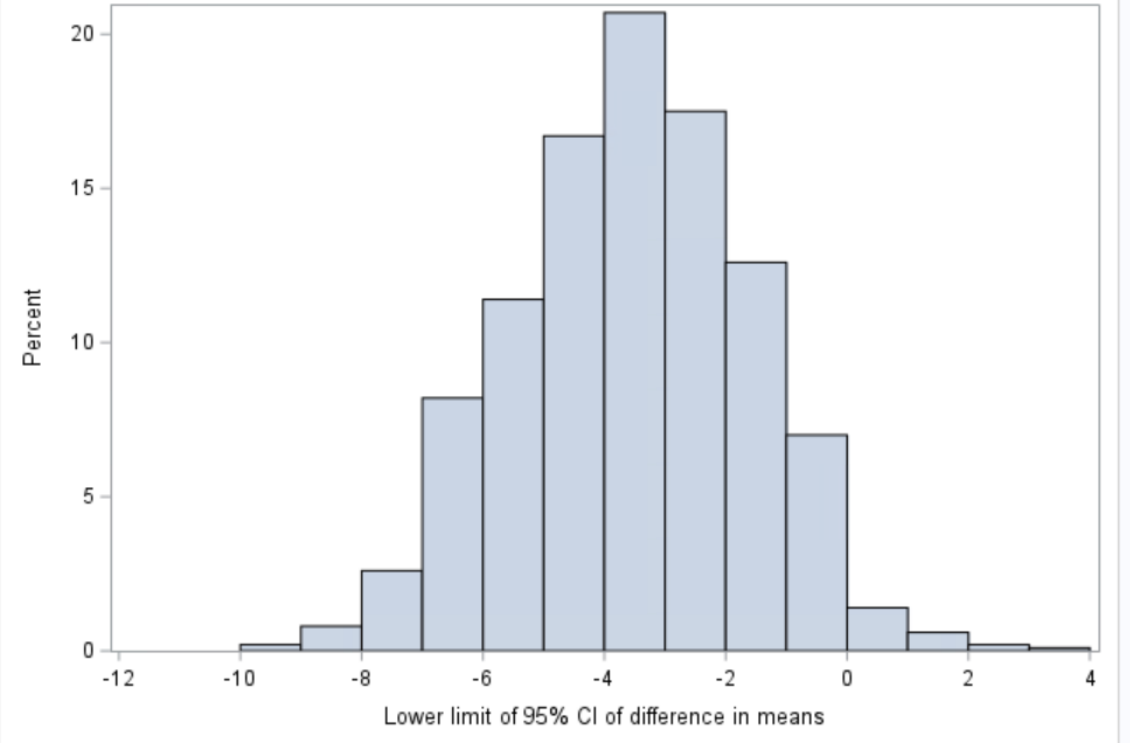


Figure 3: Histogram of Lower Limit of 95% Confidence Interval for Difference in Mean Reduction in % Glucose Values <70 mg/dL Between DTX401 and Placebo



Derivation of NI margin for the endpoint of "change from baseline in percent of glucose values in 70-120 mg/dl range"

GSDIa patients who do not use cornstrach will have close to 0% glucose values in 70-120 mg/dL range as they would experience severe hypoglycemia without the use of cornstrach. Based on historical data (described below), the % of glucose values in 70-120 mg/dL range for GSDIa patients using cornstrach is ~ 75% , hence M1 (difference in % of glucose values in 70-120 mg/dL range between patients using cornstarch vs patients not using cornstarch) will be very large (~75%) and M2 (which is 50% of M1) will be large too. Hence, as described earlier, clinical judegment was sought from multiple KOLs to determine the clinically relevant margin. The KOLs determined that 15% difference in % of glucose values in 70-120 mg/dL would be clinically reasonable and within the threshold for clinical stability. Based on this, 15% margin was chosen for testing this endpoint.

This margin appears suitable for this study based on the precision of difference in effects between DTX401 group and Placebo group (GSDIa patients using cornstarch) expected in this study as per the simulations described below.

The mean and SD of % glucose values in 70-120 mg/dL observed in historical data (described above) in GSDIa patients using cornstarch for their disease management are included below.

% glucose values in 70-120 mg/dL	DTX401-CL001 (N=15)	401GSDIA01 (N=4)	(Rossi et al., 2022) (N=10)
Mean	75.5	73.8	Not available
SD	10.0	14.4	Not available

Based on the pooled analysis of the data from DTX401-CL001 and 401GSDIA01, the mean (SD) of the % of glucose values in 70-120 mg/dL is estimated as 75% (11%).

Using these estimates 1000 datasets were simulated, each containing % of glucose values in 70-120 mg/dL at baseline and Week 48 for 25 subjects in DTX401 group and 25 subjects in Placebo group. For both DTX401 and Placebo group, baseline and Week 48 values are simulated from bivariate normal distribution with mean at baseline as 75%, mean at Week 48 as 75%, SD at baseline as 11%, SD at Week 48 as 11%, and correlation between baseline and week 48 as 0.4. This is under the assumption that there is no difference between 2 groups at baseline and both groups maintained the same level of glycemic control as at baseline over time. The estimate of the difference in mean change from baseline to Week 48 in % of glucose values in 70-120 mg/dL and the corresponding 95% CI was calculated for each dataset using 2-sample t-test. Figure 4 shows the histogram of the difference in mean change from baseline between 2 groups based on the estimates in 1000 datasets. It shows that lower limit for the difference can be up to –12%.

Figure 5 shows the histogram of the lower 95% confidence limit for the difference in the mean change from baseline between 2 groups based on estimates in 1000 datasets. It shows that the lower limit of 95% CI for the difference can be up to –18%.

This indicates that even if no difference is assumed between 2 groups, because of the variability in the endpoint data, the lower limit of the confidence interval of the difference between treatment groups can be up to -18%. Based on this, 15% is chosen as a margin for testing non-inferiority (NI) of DTX401 to Placebo for this endpoint under the assumption of SD as 11%. If

the study analysis results show much larger SD, showing statistically compelling evidence of non-inferiority may not be feasible. In that case demonstration that the reduction in cornstarch intake in the DTX401 group is not accompanied by any deterioration in glycemic control will be based not only on this non-inferiority test, but on the totality of information generated in respect of changes in glucose.

Figure 4: Histogram of Difference in Mean Reduction in % Glucose Values 70-120 mg/dL Between DTX401 and Placebo

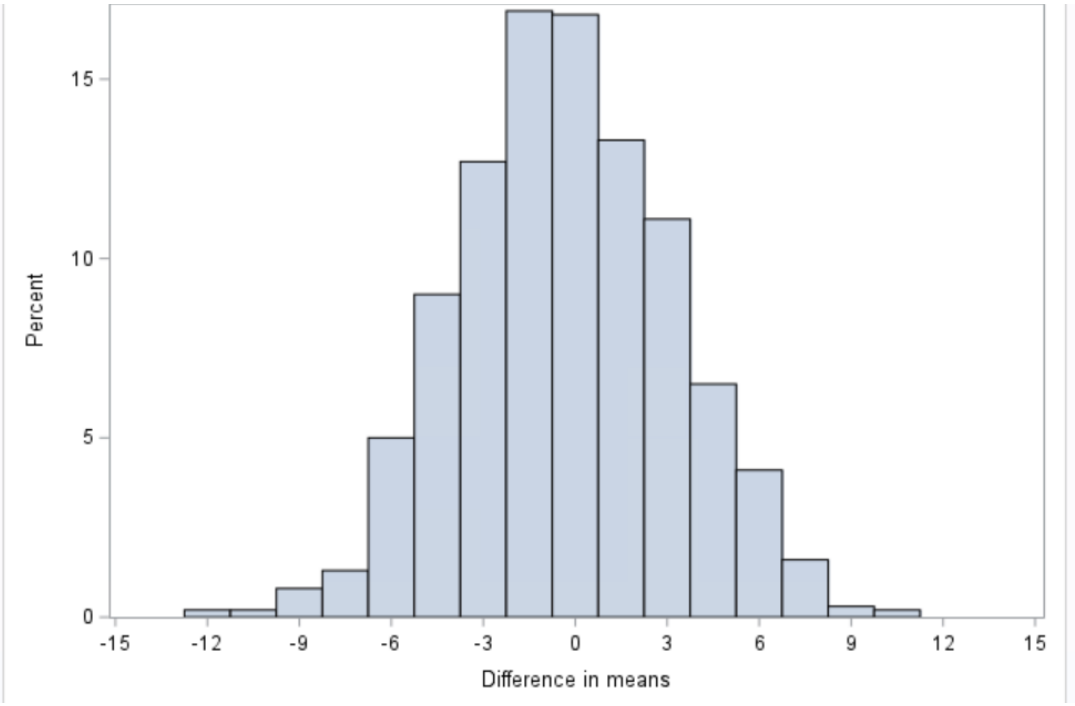
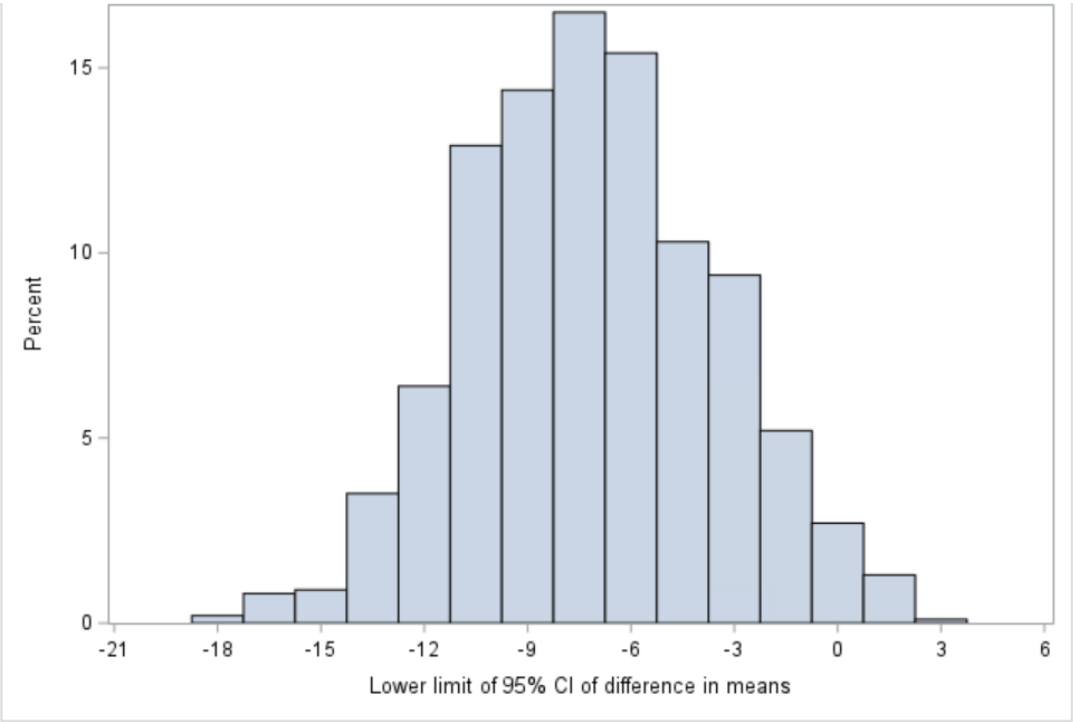


Figure 5: Histogram of Lower Limit of 95% Confidence Interval for Difference in Mean Reduction in % Glucose Values 70-120 mg/dL Between DTX401 and Placebo



APPENDIX 6. DETERMINATION OF INTERCURRENT EVENTS

Potential ICEs impacting primary and secondary efficacy endpoints are as follows:

1. ICE1: Transient changes in dietary management or nutritional intake due to acute events, such as trauma, surgery, infection, metabolic decompensation, etc.
2. ICE2: Persistent changes in dietary management or nutritional intake due to intentional or foreseeable physiologic changes, such as seasonal changes in activity, a new weight loss regimen, exercise program, puberty, pregnancy, etc.
3. ICE3: Sub-acute changes in glucose homeostasis due to adjustments in corticosteroid dose or other medications affecting glucose homeostasis, such as insulin.
4. ICE4: Partial dosing i.e., IP infusion stopped before completion and was not resumed due to adverse event (e.g., infusion related reaction) or technical issues.

Process for determining ICEs:

For ICE1 and ICE2, significant changes in cornstarch intake (grams) or carbohydrates intake (grams) from non-cornstarch diet during the period of 4 weeks prior to Weeks 24, 30, 36, 42, and 48 visit will be identified. Significant changes in cornstarch intake are defined as more than 10% change in prescribed daily cornstarch intake from the previous prescription. Significant changes in carbohydrates from non-cornstarch diet are defined as more than 25% change in prescribed daily carbohydrates (non-cornstarch diet) from the previous prescription. Thresholds of 10% and 25% are consistent with thresholds in the study protocol for cornstarch or diet modification and determination of stability during screening.

- ICE1: Transient changes due to acute events will be identified by reviewing adverse events with onset/ongoing in the 6 weeks prior to the visit. If there are no such events, these will not be considered as ICEs. If there are such AEs, site will be consulted to assess whether changes are made due to these events and accordingly these events will be documented as ICEs.
- ICE2: Persistent changes due to change in subject's needs will be identified by reviewing the reason for modification of cornstarch and diet prescription as documented on eCRF. If the reason for modification is for managing glycemic control, it will not be considered as an ICE. For other reasons, the site will be consulted to assess if changes were made due to change in subject's needs and accordingly these events will be documented as ICEs.

For ICE3, significant instability in glycemic control assessed via CGM during 4 weeks prior to Weeks 24, 30, 36, 42, and 48 visit will be identified in subjects who may have received prednisolone/placebo or concomitant medications affecting glucose homeostasis any time between Week 18 through Week 48.

Following is the list of medications that may affect glucose homeostasis and are identified based on blinded review of concomitant medications used in this study.

- ATC classification level 2 as "Corticosteroid for systemic use"
- Preferred term as Glucose and ATC classification level 2 as "General Nutrients"

- Preferred term as Prednisone, Betamethasone; Gentamicin, Dexamethasone, Steroids, Glucomannan, Insulin Glargine

The site investigator will be consulted to assess if glycemic control was impacted due to these medications. Significant instability will be defined as more than 15% variation in weekly percentage of CGM values in 70-120 mg/dL over the 4 weeks. A threshold of 15% is consistent with threshold used for assessing glycemic stability prior to dosing.

For ICE4, partial dosing at Day 1 is identified as interruption of infusion on Day 1 without restarting. These will be included as PDs.

Document Name: DTX401-CL301 Statistical Analysis Plan Version 1.0 (05APR2024)
Version #: 6.0

Workflow Step Name: Approval	Approval Signature and Date Signee Name: PPD Signature Capacity: PPD Signature Date: 09-Apr-2024 03:25:09 GMT+0000

Workflow Step Name: Approval	Approval Signature and Date Signee Name: PPD Signature Capacity: PPD Signature Date: 09-Apr-2024 17:21:16 GMT+0000

Workflow Step Name: Approval	Approval Signature and Date Signee Name: PPD Signature Capacity: PPD Signature Date: 12-Apr-2024 03:35:10 GMT+0000

Workflow Step Name: Approval	Approval Signature and Date Signee Name: PPD Signature Capacity: PPD Signature Date: 12-Apr-2024 13:52:35 GMT+0000

Document Number: VV-CLIN-002904
Global UUID: 15000_107708