Statistical Analysis Plan

4D Pharma Plc

BHT-II-002

A PHASE 2 RANDOMISED, DOUBLE BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP, MULTICENTRE STUDY TO EVALUATE THE SAFETY AND EFFICACY OF REPEATED ORAL DOSES OF Blautix™ IN ADULT SUBJECTS WITH IRRITABLE BOWEL SYNDROME (IBS) SUBTYPES IBS-C and IBS-D

Protocol Version: 31Mar2020 v3.1

Sponsor: 4D Pharma Plc
9 Bond Court
Leeds, U.K.
LS1 2JZ

Prepared by: PPD
[redacted]

Version	Date
1.0	26FEB2020
2.0	16SEP2020

[redacted]

Approval

Upon review of this document, including the table, listing, and figure shells, the undersigned approves the statistical analysis plan. The analysis methods and data presentation are acceptable.

Signature	Date
PPD	PPD
PPD 4D Pharma PLC	PPD
PPD 4D Pharma PLC	PPD

TABLE OF CONTENTS

LIST OF ABBREVIATIONS 5
1. INTRODUCTION 6
2. STUDY DOCUMENTS 6
3. STUDY OBJECTIVES 6
3.1 PRIMARY OBJECTIVE 6
3.2 SECONDARY OBJECTIVE 6
3.3 EXPLORATORY OBJECTIVES 6
4. STUDY DESIGN AND PLAN 7
5. DETERMINATION OF SAMPLE SIZE 11
6. GENERAL ANALYSIS CONSIDERATIONS 11
7. NOTATION OF TREATMENT GROUPS AND VISITS 12
7.1 NOTATION OF TREATMENT GROUPS 12
7.2 VISIT TERMINOLOGY 12
8. ANALYSIS SETS 12
9. STUDY POPULATION 13
9.1 SUBJECT DISPOSITION 13
9.2 PROTOCOL DEVIATIONS 13
9.3 DEMOGRAPHIC AND BASELINE CHARACTERISTICS 15
9.4 MEDICAL HISTORY 15
9.5 PRIOR AND CONCOMITANT MEDICATIONS 15
9.6 CONCOMITANT PROCEDURES 16
10. EFFICACY ANALYSES 16
10.1 EFFICACY ENDPOINTS 16
10.1.1 Primary Efficacy Endpoint 16
10.1.2 Secondary Efficacy Endpoints 17
10.1.3 Exploratory Efficacy Endpoints 19
10.2 BASELINE VALUES FOR EFFICACY ASSESSMENTS 19
10.3 ADJUSTMENTS FOR COVARIATES 20
10.4 HANDLING OF DROPOUTS OR MISSING DATA 20
10.5 INTERIM ANALYSIS AND DATA MONITORING 20
10.6 EXAMINATION OF SUBGROUPS 21
10.7 MULTIPLE COMPARISON/MULTIPLICITY 21
10.8 MULTICENTER STUDIES 21
11. METHODS OF EFFICACY ANALYSIS 21
11.1 PRIMARY EFFICACY ANALYSES 21
11.2 ADDITIONAL ANALYSES OF THE PRIMARY EFFICACY ENDPOINT 22
11.3 SECONDARY EFFICACY ANALYSES 23
11.3.1 Global Subject Assessment of Relief 23
11.3.2 Stool Frequency/Consistency 23
11.3.3 IBS Quality of Life Questionnaire 23
11.3.4 IBS Symptom Severity Score 24
11.3.5 Hospital Anxiety and Depression Scale 24
11.4 EXPLORATORY ANALYSES 24
12. SAFETY ANALYSES 24

12.1 EXTENT OF EXPOSURE 24
12.2 ADVERSE EVENTS 25
12.3 CLINICAL LABORATORY EVALUATION 26
12.4 VITAL SIGNS 26
12.5 PHYSICAL EXAMINATION 26
12.6 ELECTROCARDIOGRAM 26
13. CHANGES TO PROTOCOL SPECIFIED ANALYSES 26
14. REFERENCES 27
15. APPENDICES 28
15.1 APPENDIX A: PRESENTATION OF DATA AND PROGRAMMING SPECIFICATIONS 28

LIST OF ABBREVIATIONS

Abbreviation	Full Notation
ADaM	analysis data model
AE	adverse event
AR(1)	autoregressive(1)
ARH(1)	heterogeneous autoregressive(1)
ATC	Anatomical Therapeutic Chemical
BMI	body mass index
BP	blood pressure
bpm	beats per minute
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRO	contract research organization
CS	clinically significant
CSBM	complete spontaneous bowel movement
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DSMB	Data Safety Monitoring Board
ECG	electrocardiogram
eCOA	Electronic Clinical Outcomes Assessment
eCRF	electronic case report form
FAS	full analysis set
FFQ	Food Frequency Questionnaire
HR	heart rate
IBS	irritable bowel syndrome
IBS-C	constipation predominant IBS
IBS-D	diarrohea predominant IBS
IBS-QOL	IBS-Quality of Life Questionnaire
IBS-SSS	IBS-Symptom Severity Scale
ICH	International Council for Harmonisation
HADS	Hospital Anxiety and Depression Scale
LS	least squares
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed model repeated measures
NCS	not clinically significant
PT	preferred term
QC	quality control
SAE	serious adverse event

Abbreviation	Full Notation
SAF	Safety Analysis Set
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SOC	system organ class
TEAE	treatment-emergent adverse event
TLFs	tables, listings, and figures
VAS	visual analog scale
WHO	World Health Organization

1. INTRODUCTION

This document outlines the statistical methods to be implemented during the analyses of data collected within the scope of 4D Pharma Protocol BHT-II-002. The purpose of this plan is to provide specific guidelines from which the statistical analyses will proceed. Any deviations from this plan will be documented in the clinical study report (CSR). Statistical analyses presented in this plan are consistent with International Council for Harmonisation (ICH) E9, "Statistical principles for clinical trials".

2. STUDY DOCUMENTS

The following study documents are used for the preparation of the SAP:

- Protocol version 3.1, 13Feb2019
- Annotated electronic case report form (eCRF), Version 5.0, 07May2020.

3. STUDY OBJECTIVES

3.1 Primary Objective

The primary objective of the study is to assess the efficacy of repeated twice daily doses of Blautix™ >$1x10^{10}$MPN for 8 weeks in adult subjects with either constipation predominant IBS (IBS-C) or diarrhoea predominant IBS (IBS-D).

3.2 Secondary Objective

The secondary objective of the study is to assess the safety of repeated twice daily doses of Blautix™ >$1x10^{10}$MPN for 8 weeks in adult subjects with either IBS-C or IBS-D.

3.3 Exploratory Objectives

Exploratory objectives of the study are:

[redacted]

4. STUDY DESIGN AND PLAN

This is a randomised, placebo-controlled, parallel group, repeat dose multicentre study in adult male and female subjects of 18 to 70 years of age with a confirmed diagnosis of IBS-C or IBS-D.

The study consists of a Screening Visit, a Baseline Visit (Day 1), an 8-week treatment period with 2 clinic visits, Visit 2 (Week 4) and End of Treatment (Week 8), and a Follow-up Visit 4-6 weeks after the last dose of study medication for a safety check and clinical update.

The Screening Visit is up to 4 weeks before dosing starts. Subjects who fulfill all the eligibility criteria will be provided with an electronic device containing an Electronic Clinical Outcomes Assessment (eCOA) App and will be provided with training and instructions on how to use the device. At this visit, they will complete the IBS-Symptom Severity Score (IBS-SSS) on a device during the clinic visit.

The subjects will be asked to take the device home and complete the following assessments daily over a minimum of a 7-day period prior to attending the clinic for their Baseline Visit 1:

- An abdominal pain intensity rating scale daily
- Bowel movements, Stool consistency/frequency daily and laxative use; if no bowel movements, this should be recorded as zero

The subjects will need to ensure they complete these two assessments at least one day prior to the day of the next clinic visit. The same device will be used throughout the study for all patient reported outcome assessments.

Each subject will be provided with a kit with instructions for the collection of a stool sample. They will be asked to provide a stool sample from the night before or on the morning of the next clinic visit. If necessary, a stool sample may be provided up to 2 days after the Baseline visit provided it is from the night before or from the day it is delivered to the clinic.

At the Baseline visit, the study doctor will review the screening data obtained from the device eCOA app and the subject will be classified into either IBS-C or IBS-D subtypes. Subjects will be randomised to receive either Blautix™ or placebo capsules, with instructions for two (2) capsules to be taken at approximately 12-hour intervals; morning and evening, approximately 30 minutes before meals, giving a total of four (4) capsules per day.

[redacted]

Specific assessments performed at each visit are tabulated in the following Schedule of Assessments and Procedures.

SCHEDULE OF ASSESSMENTS AND PROCEDURES

	Screening Visit (Day-28 to Day-1)	Baseline Visit 1 (Day 1) Week 1	Visit 2: Week 4 (±7 days)	End of Treatment Week 8 (-7 days)	Follow-Up: Week 12-14
General Assessments					
Informed Consent (prior to any screening procedures)	x				
Inclusion/Exclusion	x	x			
Medical history/demographics	x				
Physical examination	x	x		x	x
Vital signs	x	x		x	x
Weight/Height/BMI	x	x[e]		x[e]	x[e]
ECG	x			x	
Concomitant medication	x	x	x	x	x
Adverse Event Recording [a]	x	x	x	x	x
Randomisation (8 Week Treatment Period)					
Study Drug		x	x		
Dosing Diary[b]		x	x		
Laboratory Assessments					
Blood samples (haematology and clinical chemistry)	x			x	x
Blood sample (viral serology)	x				
CCI [redacted]		x	x	x	x
Stool sample		x	x	x	x
CCI [redacted]		x	x	x	x

[redacted]

Pregnancy test [c]	x	x	x	x	x
Clinical Outcome Assessments - eCOA					
Abdominal Pain score[d]	x	x	x	x	x
Stool frequency and consistency score[d]	x	x	x	x	x
Food Frequency Questionnaire		x	x	x	x
Subject Global Assessment of Relief		x	x	x	x
IBS Quality of Life (QOL)		x	x	x	x
IBS-Symptom Severity Score (SSS)	x	x	x	x	x
Hospital Anxiety and Depression score (HADS)		x	x	x	x

[a] **Screening Visit:** Any AE's occurring prior to randomisation will be recorded as medical history; any SAEs or study related AEs that occur after signing the ICF but before randomisation will be recorded

[b] **Dosing Diary**: Patients will be provided with a smartphone at screening to capture Clinical Outcome Assessments (eCOA) and study treatment dosing data.

[c] **Pregnancy Test:** Serum at screening and Urine Dipstick test for all other visits. Female patients of childbearing potential are to continue using highly effective (refer to Protocol Section 4.2.1) contraception for two (2) menstrual cycles after the last dose of study medication. Male subjects, who are not vasectomised must use a barrier method of birth control from randomisation until the follow-up visit.

[d] **Abdominal Pain Score and Stool Frequency and Consistency score:** These assessments will be recorded daily from randomisation and over a 7-day timeframe prior to randomisation to allow the investigator to classify into IBS-D or IBS-C subtypes

[e] **Weight** only to be measured and BMI calculated, after the screening visit

5. DETERMINATION OF SAMPLE SIZE

Within each cohort (IBS-C or IBS-D) 125 subjects per treatment group (Blautix™ or Placebo) will provide at least 80% power to demonstrate a statistically significant difference in response rate between treatment groups at a one-sided alpha of 10% using Pearson's test with Yates' correction when the true response rate in the Placebo group is 40% and the true response rate in the Blautix™ group is 55%.

6. GENERAL ANALYSIS CONSIDERATIONS

The statistical analyses will be reported using summary tables, listings, and figures (TLFs). The International Council for Harmonisation (ICH) numbering convention will be used for all TLFs. Statistical testing of the primary efficacy endpoint will be one-sided and will be performed at the 0.10 significance level. For the primary efficacy endpoint, two-sided 80% confidence intervals (CI) will be presented and the 95% confidence interval will be presented as an exploratory analysis. Unless stated otherwise, statistical testing of other efficacy endpoints will be two-sided and will be performed at the 0.05 significance level, and 95% CIs will be presented. Differences in means produced by t-tests will be calculated using the Satterthwaite method.

Continuous variables will be summarized with means, standard deviations, medians, minimums, and maximums.

Categorical variables will be summarized by frequency and by percentage of subjects in corresponding categories. Percentages for missing values are omitted and do not account for the percent calculation of other categories. Percentages are routinely based on the total category count excluding the missing category if not otherwise mentioned. Footnotes will specify the percent basis in other cases.

Generally, summary tables will be presented by treatment group within each of the IBS subtypes (IBS-C and IBS-D). As appropriate, treatment group summaries and analyses will be done by combining the subtypes. Demographic, baseline summaries, and prior/concomitant medication summaries will additionally include summary columns with totals across the subtype groups for each subtype and overall.

Individual subject data obtained from the eCRFs, external vendors, the eCOA system, central clinical laboratory, electrocardiogram (ECG), and any derived data will be presented by IBS subtype, treatment, and subject in data listings.

The analyses described in this plan are considered a priori, in that they have been defined prior to database lock and breaking the study blind.

Any analyses performed subsequent to breaking the blind will be considered post hoc and exploratory. Post hoc analyses will be labeled as such on the output and identified in the CSR.

All analyses and tabulations will be performed using SAS® software Version 9.4 or higher. Tables, listings, and figures will be presented in RTF format.

The process for SAS program validation and quality control (QC) for programs and outputs involves double independent programming of tables and datasets. Listings and figures are code reviewed. All output delivered to the Sponsor undergoes review by the Lead Biostatistician and by a senior reviewer.

7. NOTATION OF TREATMENT GROUPS AND VISITS

7.1 Notation of Treatment Groups

The following notation of subtypes and treatment groups will be used throughout the report:

Subtype	***Treatment, as used throughout all tables, listings and figures***
IBS-C IBS-D IBS-C and IBS-D	Blautix Placebo

7.2 Visit Terminology

Analysis and study visits

Study day is measured from date of first dose of Blautix™ or Placebo. For presentations of data over time at each scheduled visit, Unscheduled visits are not used in summaries produced by study visit. These will be included in a summary of a derived "minimum" and "maximum" ever observed, to capture extreme values that may trigger and unscheduled visit, that is not included

For IBS Quality of Life (IBS-QOL) and Hospital Anxiety and Depression Scale (HADS) efficacy assessments by the patient on the eCOA device will have study visits assigned based on the date of the in-clinic stool sample (which has the same schedule of assessments), a window of+/- 7 days will be used, if a IBS-QOL or HADS assessment does not fall into a visit window then the assessment will be assigned as an unscheduled visit, if multiple assessments falls into the same visit window then the assessment closest to the in-clinic assessment will be selected as the scheduled visit and the other assessment will be assigned as an unscheduled visits.

For other efficacy assessments by the patient on the eCOA device (eg, abdominal pain intensity, stool consistency/frequency), weekly visits will be obtained by using 7-day intervals up to Day 56 (Week 1: 1-7 days, Week 2: 8-14 days, Week 3: 15-28 days, …, Week 8: 50-56 days). Additional data will be summarized at follow-up visits from 6-8 weeks after EOT (Week 8) using 3 created time points: Follow-up (Week 12): 78-84 days, Follow-up (Week 13): 85-91 days, and Follow-up (Week 14): 92-98 days.

Otherwise, eCRF visit will be used and study day is defined as follows:

- Assessment date - date of first dose + 1, if assessment date is on or after the date of the first dose.
- Assessment date - date of first dose, if assessment date is before the date of the first dose.

8. ANALYSIS SETS

The Safety Analysis Set (SAF) includes all subjects randomised into the study who received at least one dose of Blautix™ or Placebo. Subjects will be allocated to the treatment group corresponding to the treatment received. As we do not include the kits dispensed to patients in the clinical database, actual treatment is equal to randomised (planned) treatment. [this will be reassessed prior to database lock]

The Full Analysis Set (FAS) includes all subjects in the SAF who were appropriately randomised [This will be defined prior to database lock] into the study. Subjects will be allocated to the treatment group corresponding to the to the actual treatment received. The FAS is the primary analysis set for the efficacy analyses. Subsets of subjects in the FAS who have valid baseline data for assessments of change in relation to response and improvement endpoints will be used for the corresponding analyses.

The Efficacy Evaluable Set includes all members of the Full Analysis Set who completed the 8-week treatment and assessment period (have a visit between 49 and 64 days) without major protocol violations deemed to impact the assessment of efficacy. Subjects will be allocated to the treatment group corresponding to the actual treatment received. The Efficacy Evaluable Set will be used for sensitivity analyses of the primary and secondary efficacy endpoints.

The final determination of the membership of analysis sets will be made at a blinded data review meeting before database lock convened by the Sponsor.

8.1 Cohort Definition

All outputs including IBS subtype will be based on the value entered into the IVRS system, which cannot be missing.

9. STUDY POPULATION

9.1 Subject Disposition

Subject disposition information will be summarized for all subjects by IBS subtype and treatment group. Summary columns will also be included for subtype and overall totals. Summaries will include: the number of subjects screened and the number of screen failures, the number of subjects randomised, the number and percentage of subjects in each analysis set, the number and percentage of subjects completing the study, and the primary reason for discontinuation. The number of subjects randomised will be the denominator for percentages.

A separate summary will show the reasons for screen failure with percentage of subjects who failed screening for each reason. The reasons for screen failure will also be listed.

9.2 Protocol Deviations

Protocol deviations are collected in the Parexel and Synteract Clinical Trial Management Systems, which will be pooled. Deviations are reviewed monthly and classified as major or minor. Major protocol deviations that could potentially affect the efficacy of the study will be identified prior to database lock and unblinding of individual subject treatment information. This group of major deviations will be used to define the EE population. Major protocol deviations may include, but are not limited to:

- Randomly assigned subjects who did not satisfy selected inclusion and exclusion criteria;
- Randomly assigned subjects who developed withdrawal criteria during the study but were not withdrawn;
- Subjects who received the wrong treatment or incorrect dose;

- Subjects who received a prohibited concomitant medication as noted below:

IBS-C Patients:

Guanylate cyclase agonists:

Linaclotide

Lubiprostone

IBS-D PATIENTS:

5-hydroxytryptamine (serotonin) 3 receptor antagonists:

Alosetron

Bile Acid sequestrants:

Cholestyramine

Colesevelam

Colestipol

Eluxadoline

Rifaximin

PROBIOTICS:

Alflorex (EU)
Align (North America)
Align ® (B. longu m infantis 35 624) Capsules
Align Probiotic Supplement Capsules ((Lactobacillus GG)
BioGaia L. Reuteri ProTectis Probiotic (Lactobacillus) Capsules
Bioglan products
Bio-Kult
Culturelle Probiotic (Lactobacillus GG) Capsules
Digestive Care™ (L. plantarum 299v) Capsules
Human MicroFlora products by Genestra™
IbSium® (Saccharomyces cerevisiae I-3856) Capsules
Mutaflor ® (active ingredient Escherichia coli strain Nissle 1917)
Prescript-Assist (Broad Spectrum Probiotic Prebiotic) Capsules
ProBio 7 products
Proxiflor ® (L. rhamnosus R0011 L.helveticus R0052)Capsules
TuZen® (L. plantar um 299v) Capsules
Udos probiotics
Ultra Probiotic Complex by GNC-

UltraFlora Intensive Care (L. planta rum 299v) Capsules
Vitabiotic products
Yakult (Lactobacillus casei Shirota) packet

The decision whether a subject is excluded from the Efficacy Evaluable Analysis Sets will be made during the data review meeting. Reasons for exclusion of a subject from the analysis sets, and protocol deviations (minor/major) will be listed.

All protocol deviations will be summarized by deviation category (all, major, minor), IBS subtype, and treatment group.

9.3 Demographic and Baseline Characteristics

Demographic variables include: age, sex, ethnicity, and race. Age is provided by the IVR system and not derived.

Other characteristics include: height, weight, and body mass index (BMI).

Baseline disease characteristics to be summarized are the weekly average Abdominal Pain Intensity, weekly average stool frequency, and weekly proportion (as %) of days with at least one stool of Type 6 or 7 on the Bristol Stool Chart.

Descriptive statistics will be presented for age, height, weight, BMI, and other continuous variables. Frequency counts and percentages will be presented for sex, ethnicity and race. Age will also be summarized by groups (≤ 44 years, 45 – 64 years, and ≥ 65 years). The denominator for percentages will be the number of subjects in the IBS subtype and treatment for the analysis set. Demographic variables and characteristics will be summarized for the Safety, Full, and the Efficacy Evaluable Analysis Sets.

9.4 Medical History

Medical History verbatim terms in the eCRFs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Medical History will be summarized by primary MedDRA System Organ Class (SOC) and preferred term (PT) within the SOC. Subjects will be counted once at each level.

9.5 Prior and Concomitant Medications

Prior and concomitant medication verbatim terms in the eCRFs will be mapped to Anatomical/Therapeutic/Chemical (ATC) class and Preferred Names using the WHO Drug Dictionary (WHODD) Global.

Prior medications are those medications that started before the initial dose of study drug and end prior to the first dose. Concomitant medications are those medications that started on or after the date of the initial dose of study drug or medications that started before the initial dose of study drug and continued during the treatment period.

Prior and concomitant medications will be summarized for each treatment by WHO ATC class level 3 and preferred name. These summaries will present the number and percentage of subjects using each medication.

Subjects may have more than 1 medication per ATC class and preferred name. At each level of subject summarization, a subject is counted once if he/she reported 1 or more medications at that level. Each summary will be ordered by descending order of incidence of ATC class and descending incidence of preferred name within each ATC class as observed in all subjects. Ties will be broken alphabetically.

Changes to concomitant medications, including new medications and stopping of current medications are recorded in the eCOA device. This data will be entered in the eCRF.

9.6 Concomitant Procedures

Concomitant procedures recorded in the eCRF will be listed only.

10. EFFICACY ANALYSES

The primary efficacy analysis will be based on the FAS. Additional supportive efficacy analyses will be performed using the Efficacy Evaluable Set.

10.1 Efficacy Endpoints

The efficacy endpoints in this study are derived from daily eCOA diary entries and from questionnaires administered during the Week 4 and Week 8 (EOT) visits.

10.1.1 Primary Efficacy Endpoint

The primary efficacy endpoint is whether the subject is an overall responder. The definition of an overall responder is IBS subtype specific and consists of two components.

For both IBS subtypes, the first component of overall response is the weekly average of Abdominal Pain Intensity (0-10, 0=none, 10=worst possible pain) over the past 24 hours. Improvement for the week is defined as a decrease of at least 30% compared with baseline.

For the IBS-C subtype, the second component of overall response is the stool frequency per week. Improvement is defined as an increase of 1 or more complete spontaneous bowel movements (CSBMs) per week compared with baseline. The number of bowel movements is recorded once per day in the eCOA device, including zeroes.

For the IBS-D subtype, the second component of overall response is the proportion of days per week with at least one stool with a consistency of Type 6 or 7 on the Bristol Stool Chart. Improvement is defined as a decrease in the proportion of at least 50% compared with the proportion during the 7-day period before baseline used to determine eligibility. Stool consistency is recorded for each bowel movement in the eCOA device, so more than one entry per day may exist. The daily result should be reduced to whether a Type 6 or Type 7 was included on that date.

A subject will be considered a responder only during weeks for which both components indicate improvement. Subjects must have at least 4 evaluable weeks of data as defined in Section 10.4 for each of the instruments used to determine response. Subjects with less than 4 evaluable weeks of assessments available during the treatment period will be considered as non-responders for overall response.

Subjects who have at least 7 evaluable weeks of data and have reported an improvement in their weekly symptoms for ≥ 50% of the treatment period will be considered overall responders. Subjects who have 4, 5, or 6 evaluable weeks of data must show improvement for ≥ 60% of the treatment period to be deemed overall responders.

The days for each week will be taken as calendar days from the Day 1 date (randomisation date).

The denominators for the weekly average of Abdominal Pain Intensity, the stool frequency, and the proportion of days per week with at least one Type 6 or 7 will be the number of days with an assessment during the week. The stool frequency will be multiplied by 7 to convert to a stools per week figure. The stool consistency will be multiplied by 100 to convert to a percentage.

10.1.2 Secondary Efficacy Endpoints

10.1.2.1 Subject Global Assessment of Relief

The Subject Global Assessment of Relief is collected weekly through the eCOA system. It is a comparison of how the subject has felt over the past week with regards to their IBS to the way they felt before entering the study. It is measured on a 5-point Likert scale with the following responses:

Completely relieved; considerably relieved; somewhat relieved; unchanged; worse.

10.1.2.2 Stool Consistency/Frequency

Frequency of bowel movements is captured daily by the eCOA system. The consistency of each bowel movement is rated according to images in the 7-level Bristol Stool Chart. See Appendix B for a graphic representation of the chart.

Derivation of the stool frequency per week and the proportion of days per week with an increase of at least one stool with consistency of Type 6 or 7 from baseline, as described as components of the primary efficacy endpoint in Section 10.1.1.

10.1.2.3 IBS Quality of Life Questionnaire

The IBS Quality of Life (IBS-QOL) questionnaire consists of a series of 34 questions concerning the subject's feelings about their IBS over the past month. The responses to each question are on an ordinal scale and numbered 1-5 with:

1 = NOT AT ALL
2 = SLIGHTLY
3 = MODERATELY
4 = QUITE A BIT
5 = EXTREMELY or A GREAT DEAL

The possible raw score range for the numerator of the total quantity is 136. The formula transforms the item scores into a 4 to 0 range so that higher values of the total score are associated with increased quality of life. It is possible to derive eight subscale scores as well as a single global score. Each item is taken to contribute equally to each subscale with each subscales containing the following items:

Subscale	IBS-QOL Question Numbers	Possible Raw Score Range
Dysphoria	01, 06, 07, 09, 10, 13, 16, 30	32
Interference with Activity	03, 18, 19, 22, 27, 29, 31	28
Body Image	05, 21, 25, 26	16
Health Worry	04, 15, 32	12
Food Avoidance	11, 23, 28	12
Social Reaction	02, 14, 17, 34	16
Sexual	12, 20	8
Relationships	08, 24, 33	12

Subscales are scored through simple summative scaling. All items are negatively framed with the greatest response scale equaling the worst quality of life. When scored, all items are reversed so that the as IBS-QOL scores increase, quality of life increases. All final raw scores are transformed to a 0 to 100 scale using the following formula: $\frac{Sum\ of\ the\ items - lowest\ possible\ score}{Possible\ raw\ score\ range} * 100$

This transformation converts the lowest and highest possible scores to zero and 100, respectively. Scores between these values represent the percentage of the total possible score achieved. The IBS-QOL instrument and scoring programs have used this transformation to provide comparative data for interpretation.

10.1.2.4 IBS Symptom Severity Score

The IBS Symptom Severity Score (IBS-SSS) measures the severity of IBS symptoms through a series of questions and visual analog scales (VAS). Part 1 of the IBS-SSS generates the severity score, as described in Appendix 4 of the protocol. Part 2 consists of Bowel Habit and Site of Pain sub-sections that are not converted to a score.

10.1.2.5 Hospital Anxiety and Depression Scale

The Hospital Anxiety and Depression Scale (HADS) is commonly used by doctors to determine the levels of anxiety and depression that a person is experiencing. The scale consists of a series of statements about which subjects must respond with their feelings. Each statement has a set of ordinal responses with scores 0 to 3 or 3 to 0.

Some statements correspond with anxiety and some with depression, as given in the table below.

Statement	Characteristic
I feel tense or wound up	Anxiety
I get a sort of anxious feeling like butterflies in the stomach	Anxiety
I get a sort of frightened feeling as if something awful is about to happen	Anxiety
I feel restless as if I have to be on the move	Anxiety
Worrying thoughts go through my mind	Anxiety
I get sudden feelings of panic	Anxiety
I can sit at ease and feel relaxed	Anxiety
I feel as if I am slowed down	Depression
I enjoy the things I used to enjoy	Depression
I have lost interest in my appearance	Depression
I can laugh and see the funny side of things	Depression
I look forward with enjoyment to things	Depression
I feel cheerful	Depression
I can enjoy a good book, radio or television program	Depression

An Anxiety and a Depression Total Score ranging from 0 to 21 each are derived by summing the individual scores under each category. The Total Scores are interpreted as:

0-7 = Normal,
8-10 = Borderline abnormal (borderline case)
11-21 = Abnormal (case)

10.1.3 Exploratory Efficacy Endpoints

CCI [redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted]

10.2 Baseline Values for Efficacy Assessments

For the daily assessments collected on the eCOA device, baseline values will be derived from the 7 consecutive days assessed preceding the Baseline (Day 1) visit used to assess eligibility. The baseline value will be considered missing if less than 4 of the 7 days have an assessment.

- Abdominal Pain Intensity (1-10)

[redacted]

- Stool Consistency (day with at least one bowel movement Type 6 or 7)
- Stool Frequency

The baseline value for each of these assessments will be calculated by summing the daily result over the 7-day period and dividing by the number of days with an assessment. The Stool Consistency baseline proportions will be multiplied by 100 to convert to a percentage.

Baseline for the other efficacy assessments is defined as the last non-missing value recorded on or before the Baseline (Day 1) visit.

10.3 Adjustments for Covariates

The primary analysis (proportion of overall responders) will not be stratified or adjusted for any covariates.

The proportion of subjects who are overall responders will be analyzed with both subtypes combined in a stratified analysis using the Cochran-Mantel-Haenszel (CMH) statistic.

The proportion of subjects who are overall responders will be further analyzed using separate logistic regression models for each subtype including terms for treatment group, region (US or UK and Ireland), the baseline week average of Abdominal Pain Intensity, and the subtype-specific baseline value. For the IBS-D model, the baseline covariate will be the proportion (as %) of days during the 7 consecutive days prior to baseline used for eligibility, with at least one bowel movement of Type 6 or 7. For the IBS-C model, the baseline covariate will be the baseline week average stool frequency.

10.4 Handling of Dropouts or Missing Data

In general, missing data will not be imputed. Summary data for observed cases only will be presented.

For weekly assessments derived from daily eCOA data, the weekly values will be considered missing if there are less than 4 days with data assessed within a week.

For the primary endpoint, overall response, subjects must have at least 4 evaluable weeks of data. An evaluable week is one in which there are at least four 24-hour daily records of abdominal pain score, number of CSBMs, and Bristol stool chart. If number of CSBMs for a date is 0, then Bristol stool chart data is expected to be missing. The day will still count towards the 4-day minimum. Subjects with less than 4 evaluable weeks of data will be considered non-responders.

Mixed model repeated measures (MMRM), which provide valid parameter estimates even with some values missing, will be used to analyze continuous assessments taken over time.

10.5 Interim Analysis and Data Monitoring

After approximately 250 subjects have completed 8 weeks of treatment, an interim analysis will be conducted. This unblinded interim analysis will be provided to the Chief Scientific Officer (CSO) at 4D pharma. These analyses will be conducted by an unblinded study team, independent of the study team, as documented in the Unblinding Plan.

Summary tables for disposition, demographics and baseline characteristics, safety data, and other pertinent study data (e.g., concomitant medications and study drug exposure/compliance) will be constructed with treatment groups masked, yet unblinded.

The efficacy tables will consist of the table for the primary efficacy analysis and the tables covering additional analysis of the primary efficacy endpoint. The secondary analysis tables for stool consistency and frequency will also be included.

The assignment to masked treatment group will be done at random by an unblinded statistician and labeled in a way that masks treatment (e.g., Treatment A, Treatment B.).

The protocol describes an interim analysis of the unblinded efficacy data with early stopping for futility. It has been decided that the trial will not be stopped for futility.

The DSMB will not review the interim analysis. The DSMB will be presented the final study results after unblinding and database lock.

10.6 Examination of Subgroups

The summary and analysis of the primary efficacy endpoint (overall response) will be repeated for subgroups of sex and region (US, UK and Ireland).

The CMH analysis of overall response stratified by subtype will also be presented by sex and by region.

10.7 Multiple Comparison/Multiplicity

No adjustments for multiplicity will be made in this study.

10.8 Multicenter Studies

This is a multicenter study, having about 30 centers participating in the study. Only about 8 of these centers have high enrollment, with relatively low enrollment at the other centers. The low number of subjects at many centers does not allow the center to be included as a factor in statistical models. Region (US, UK and Ireland), will be used in supplemental analysis of the primary efficacy endpoint and in secondary efficacy analyses as specified.

11. METHODS OF EFFICACY ANALYSIS

11.1 Primary Efficacy Analyses

For the primary analysis of the overall response rate, as defined in Section 10.1.1, the following one-sided hypotheses will be tested with a significance level α=0.10:

$$H_0: \pi_T - \pi_P \leq 0 \leftrightarrow H_1: \pi_T - \pi_P > 0$$

where π_T and π_P are the proportions of overall responders under active treatment and placebo after 8 weeks of treatment, respectively. The hypotheses will be tested using the FAS using Pearson's chi-square test with Yates' continuity correction. As this is a one-sided chi-square test, the p-value from

the test will be divided by two if the difference is in favor of the Blautix arm ($\pi_T > \pi_P$) and divided by two then subtracted from 1 if the difference is in favor of the placebo arm ($\pi_T \leq \pi_P$).

The test will be applied separately to each IBS subtype.

Summary statistics will show the frequency of overall responders in each subtype and treatment group, the percentage of responders in each treatment group within subtype, the treatment difference in percentages within each subtype (active – placebo), and 80% Wald (asymptotic) confidence intervals, with continuity correction, for the differences in percentages.

11.2 Additional Analyses of the Primary Efficacy Endpoint

As a sensitivity analysis, the primary efficacy analysis of the overall response rate will be repeated for the Efficacy Evaluable Set.

The remaining additional summaries and analyses will be done using the FAS.

The overall response rate will be analyzed with the IBS subtypes combined. A CMH test stratified by subtype will be used to test for an overall difference between the treatment groups across the subtype strata. To correspond with the primary analysis, the significance level will be two-sided 0.20 and an 80% CI for the treatment difference will be provided.

The primary efficacy endpoint (overall response) will be summarized and analyzed using the same method as the primary analysis by subgroups formed by sex and by region (US, UK and Ireland). The CMH analysis of overall response stratified by subtype will also be presented by sex and by region.

The frequencies and percentages of responders and non-responders by week during the treatment period will be summarized by IBS subtype and treatment group. A summary for the subtypes combined will be included. No CIs or inferences will be given.

The overall response rate will be further analyzed using a logistic regression model for each subtype. The models will include treatment and region (US, UK and Ireland) as factors and the baseline week average of Abdominal Pain Intensity and the subtype-specific baseline value as covariates. For the IBS-D model, the baseline covariate will be the proportion (as %) of days during the consecutive week prior to baseline, used for assessment of eligibility assessment, with at least one bowel movement Type 6 or 7. For the IBS-C model, the baseline covariate will be the average stool frequency in the week prior to baseline as noted for IBS-D. Model based odds ratios will be reported with 95% CIs.

The weekly average Abdominal Pain Intensities, the Abdominal Pain Intensity differences (baseline – weekly average, such that positive differences indicate improvement), and the percent differences [100*(baseline – weekly average) / baseline] will be summarized by IBS subtype (including combined), treatment group, and week across the treatment period and follow-up.

The weekly average Abdominal Pain Intensity differences during the treatment period (not including follow-up) will be further analyzed using linear models for repeated measures. The models will include treatment, region, and week as main effects, treatment-by-week as an interaction term, and the baseline week Abdominal Pain Intensity as a covariate. Separate models will be built for each IBS subtype and for the subtypes combined. The IBS subtype will be included in the model for combined subtypes as a fixed effect. The covariance matrix will be unstructured. If there are convergence issues, corrected AIC will be used to

choose between the following covariance structures: Toeplitz, heterogeneous autoregressive (1) [ARH(1)] and autoregressive(1) [AR(1)]. The method of Kenward and Roger will be used to estimate denominator degrees of freedom.

Least squares (LS) means and their standard errors (SE) from the MMRM models for the weekly average Abdominal Pain Intensity differences by IBS subtype, treatment group, and week with 95% CIs will be presented. The LSM and SE of the treatment differences from placebo will also be presented with 95% CIs and p-values from t-tests for a difference from 0.

Listings of daily assessments from the eCOA data will not be provided. However, the derived weekly endpoints and response rates will be listed by IBS subtype and treatment group.

11.3 Secondary Efficacy Analyses

11.3.1 Global Subject Assessment of Relief

The Global Subject Assessment of Relief will be summarized with frequencies and percentages of subjects reporting each category by IBS subtype (including combined), treatment group within subtype, and windowed week. Missing values will not be reported so the denominator for percentages will be the number of subjects with a response for the week in the subtype/treatment group. No statistical inferences will be made.

Data will be listed by windowed week, including responses not windowed.

11.3.2 Stool Frequency/Consistency

The mean derived weekly stool frequencies, mean changes from baseline, and mean percent changes from baseline will be summarized by IBS subtype, treatment group, and week. No summary will be done with the combined subtypes. No statistical inferences will be made.

Percent change will be calculated as follows:

- If baseline value is 0, change from baseline * 100
- Otherwise, change from baseline/baseline result * 100

The number of CSMBs will be summarized by week including the changes from baseline. Since these quantities are not expected to follow a normal distribution, the median changes from baseline and median percent changes from baseline will be compared using the Mann-Whitney test.

The percentage of days each week with at least one stool with a consistency of 6 or 7 will be summarized for IBS-C and a consistency of 1 or 2 will be summarized for IBS-D using the Bristol Stool Chart.

The derived data will be listed by evaluable week.

11.3.3 IBS Quality of Life Questionnaire

The responses to each of the 34 questions on the IBS-QOL will be summarized with frequencies and percentages of subjects reporting each category by IBS subtype, including combined, treatment group within

subtype, and visit. Missing values will not be reported so the denominator for percentages will be the number of subjects with a response for the visit in the subtype/treatment group.

The mean overall total score and the 8 subscale scores and mean changes from baseline will be summarized by IBS subtype (including combined), treatment group, and visit. Means for subscales with questions missing responses will not be calculated. P-values from t-tests comparing mean change from baseline between Blautix and placebo will be presented at each post-baseline visit for exploratory purposes.

All response data from the IBS-QOL and the derived scores will be listed.

11.3.4 IBS Symptom Severity Score

The IBS-SSS mean scores from Part 1 and mean changes from baseline will be summarized by IBS subtype (including combined), treatment group, and visit. P-values from t-tests comparing mean change from baseline between Blautix and placebo will be presented at each post-baseline visit for exploratory purposes.

Analyses of the IBS-SSS will be repeated excluding data that was collected over the phone and transcribed due to COVID-19 remote visits. If notable differences are observed with and without these visits, additional analyses may be completed to properly control for this change in methodology.

All response data from the IBS-SSS will be listed (ie Part 1 and Part 2).

11.3.5 Hospital Anxiety and Depression Scale

The mean Anxiety and Depression scores for the HADS and mean changes from baseline will be summarized by IBS subtype (including combined), treatment group, and visit. P-values from t-tests comparing mean change from baseline between Blautix and placebo will be presented at each post-baseline visit for exploratory purposes. Means for subscales with questions missing responses will not be calculated.

All response data from the HADS and the derived Anxiety and Depression scores will be listed.

11.4 Exploratory Analyses

CCI [redacted]

12. SAFETY ANALYSES

All safety analyses will be based on the safety set. For summaries of laboratory values and vital signs by visit, unscheduled visits will not be presented. To capture all potential safety signals, the maximum and minimum values observed at any time during the study (including unscheduled visits) will be presented in addition to scheduled timepoints.

12.1 Extent of Exposure

The first and last dates of study drug administration are collected in the eCRF. The duration of treatment in days will be calculated as Last Dose Date - First Dose Date + 1.

Actual dosing is collected in the eCOA device as whether morning dose was taken and whether evening dose was taken. This data will be used to compute total capsules taken (2 per dose) and compliance.

Calculation of compliance requires an expected number of capsules taken. Subjects may have taken only the evening dose on the day of the Baseline (Day 1) visit. Likewise, only the morning dose may have been taken on the last day of dosing. As an approximation to the expected number of capsules taken, 4 times the duration of treatment in days minus 1 will be used. That is, compliance as a percent will be calculated as:

100 × (2 × (Sum of doses taken from eCOA)) / (4 × (Duration of treatment - 1))

Study drug exposure will be summarized by IBS subtype (including combined) and treatment group for each treatment using the duration of treatment and the total number of capsules taken.

Study drug compliance will be summarized using frequencies and percentages as categorized below:

- >110%
- >90%-110%
- >80%-90%
- >70%-80%
- <=70%

12.2 Adverse Events

All AE summaries will be restricted to treatment-emergent AEs (TEAEs), which are defined as those AEs that occurred or worsened on or after the first day of dosing. Reported AEs should also include existing AEs that worsened during the study. If it cannot be determined whether the AE is treatment emergent due to a partial onset date, then it will be counted as such. Verbatim terms in the eCRFs will be mapped to preferred terms and system organ classes using MedDRA.

Each AE summary will be displayed by IBS subtype (including combined) and treatment group. Summaries that are displayed by system organ class and preferred terms will be ordered by descending order of incidence of system organ class and preferred term within each system organ class as observed in all subjects. Ties will be broken alphabetically. Summaries of the following types will be presented:

- Overall summary of TEAEs that contain an overview of each item below.
- Subject incidence of TEAEs and total number of unique TEAEs by MedDRA system organ class and preferred term.
- Subject incidence of TEAEs by MedDRA system organ class, preferred term, and highest severity. At each level of subject summarization, a subject is classified according to the highest severity if the subject reported 1 or more events. Adverse events with missing severity will be considered severe for this summary.
- Subject incidence of TEAEs by MedDRA system organ class, preferred term, and closest relationship to study drug (Possibly Related/Not Possibly Related). At each level of subject summarization, if a subject reported 1 or more events, a possibly related TEAE will be counted over a not possibly related TEAE. Adverse events with a missing relationship will be considered related for this summary.
- Subject incidence of serious TEAEs and total number of unique serious TEAEs by MedDRA system organ class and preferred term.

- Subject incidence of TEAEs leading to death as an outcome by MedDRA system organ class and preferred term.
- Subject incidence of TEAEs leading to study discontinuation by MedDRA system organ class and preferred term.

12.3 Clinical Laboratory Evaluation

Laboratory parameters (serum chemistry and hematology) will be summarized using descriptive statistics at baseline, at each postbaseline time point, and for the minimum and maximum values observed (including unscheduled visits). Changes from baseline will also be summarized.

In addition, shift tables (i.e., low-normal-high at baseline versus low-normal-high at follow-up in a 3-by-3 contingency table) will be provided to assess changes in laboratory values from baseline to follow-up.

12.4 Vital Signs

Vital signs will be summarized using descriptive statistics at baseline, at each post baseline time point, and for the minimum and maximum values observed (including unscheduled visits). Changes from baseline will also be summarized.

12.5 Physical Examination

Since any physical examination findings will be reported as AEs, physical examination results will be included in data listings only.

12.6 Electrocardiogram

Overall interpretation results for ECG will be summarized using shift tables (Normal, Abnormal Not Clinically Significant, Abnormal Clinically Significant) comparing baseline to Week 8.

13. CHANGES TO PROTOCOL SPECIFIED ANALYSES

The EE analysis set will not be based on all major protocol deviations as indicated in the protocol, but on a subset of the major deviations that are deemed to have significant impact on efficacy data and will be determined in the data review meeting, prior to database lock and prior to unblinding the study.

14. REFERENCES

US Department of Health and Human Services (DHHS), Food and Drug Administration, Center for Drug Evaluation and Research (CDER). Guidance for Industry ICH E9 Statistical principles for clinical trials. September 1998 [cited 2018 Aug 03]. Available from: https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073137.pdf

15. APPENDICES

15.1 Appendix A: Presentation of Data and Programming Specifications

General

- Specialized text styles, such as bold, italics, borders, shading, superscripted, and subscripted text will not be used in tables, figures, and data listings unless they add significant value to the table, figure, or data listing.
- Only standard keyboard characters are to be used in tables and data listings.
- Special characters, such as nonprintable control characters, printer-specific, or font-specific characters, will not be used on a table, figure, or data listing.
- Hexadecimal character representations are allowed (e.g., μ, α, ß).
- All footnotes will be left justified and at the bottom of a page. Footnotes must be used sparingly and must add value to the table, figure, or data listing.

Tables

- Formal organization of tabulations may be changed during programming, if appropriate, e.g., tables for the different variables may be combined into a single table, or tables with more than 1 variable may be split into several tables.
- Means and medians will be presented to 1 more decimal place than the raw data. Standard deviations will be presented to 2 more decimal places than the raw data. Minimums and maximums will be reported with the same number of decimal places as the raw data.
- Percentages will be presented to the tenths place.
- For frequency counts of categorical variables, categories whose counts are zero will be displayed for the sake of completeness. For example, if none of the subjects discontinue due to "lost to follow-up," this reason will be included in the table with a count of 0. Categories with zero counts will not have zero percentages displayed.
- Lower and upper CI values must be presented to 1 decimal place more than the raw/derived data (i.e., to the same number of decimal places as the mean).
- Percentiles (e.g., 25%, 75%) must be presented to 1 decimal place more than the raw/derived data.
- For all inferential analyses, *P* values will be rounded to 3 decimal places (or at the highest level of precision) with a leading zero (0.001). *P* values less than 0.001 will be presented as "<0.001."
- The last footnotes will be
  - "PROGRAM SOURCE: ...\\xx.sas, DATA CUT OFF DATE: DDMMMYYYY, RUN DATE: DDMMYY hh:mm". where the extract date is the datestamp of the data snapshot used.

Figures

- Legends will be used for all figures with more than 1 variable or item displayed. Treatment group sizes (n=xx) will be included, as appropriate.
- Figures will be in black and white but can be in color to add value to the clarity and readability of a figure. Lines must be wide enough to see the line after being copied.

- The last footnotes will be
  - "Source: xxx", where xxx indicates the source listing number(s) and/or source datatset(s) (e.g., ADaM).
  - "PROGRAM SOURCE: ...\\xx.sas, DATA CUT OFF DATE: DDMMMYYYY, RUN DATE: DDMMYY hh:mm".

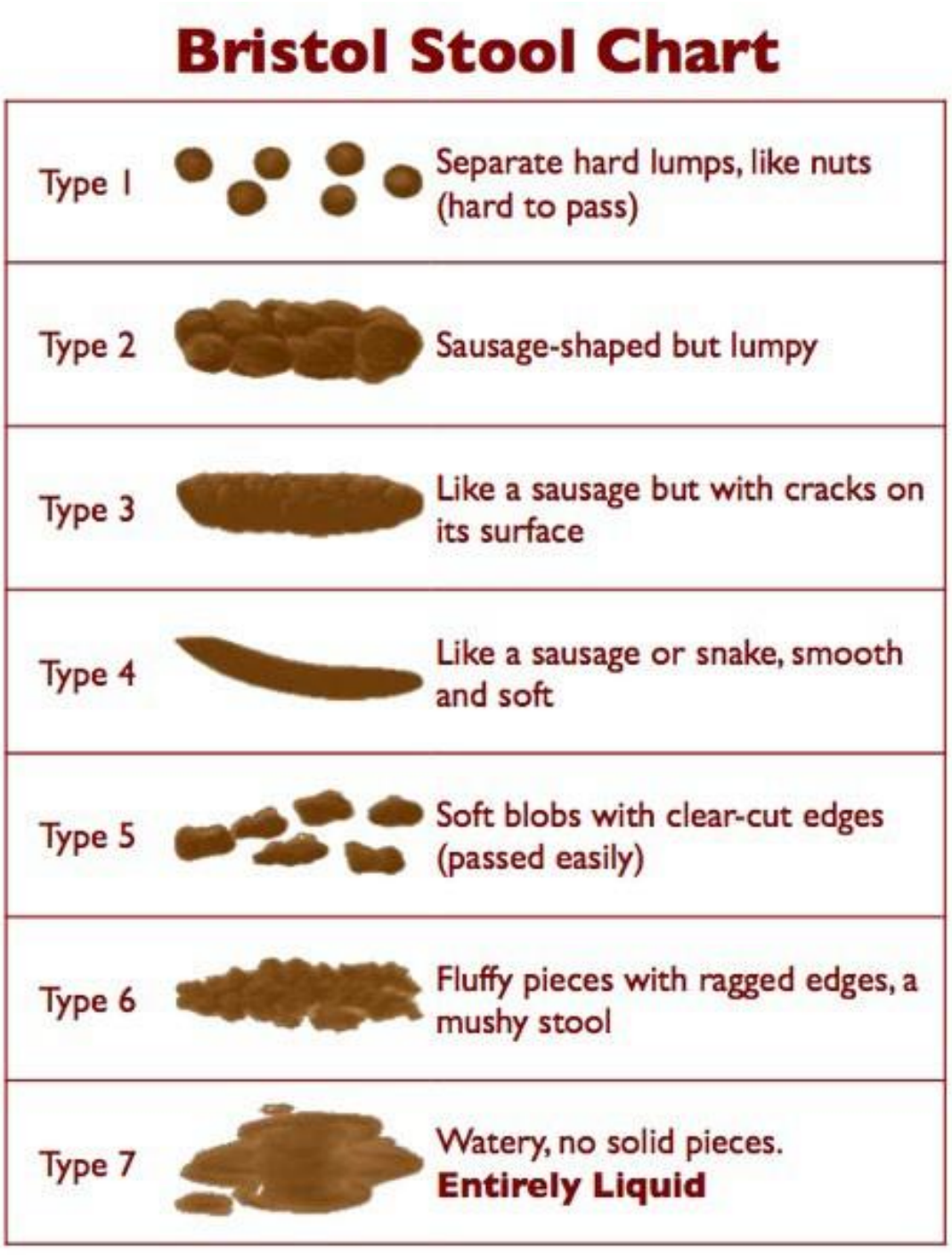
  where the extract date is the datestamp of the data snapshot used.

Listings

- Formal organization of the listing may be changed during programming, if appropriate, e.g., additional variables may be included, change in the column order, or the listing may be split into multiple parts due to space constraints.
- If not otherwise specified, all data listings will be sorted by IBS subtype, treatment, center, subject number, visit, and date/time, as appropriate.
- All date values will be presented in a SAS date (e.g., 29AUG2001) format.
- All observed time values will be presented using a 24-hour clock HH:MM:SS format (e.g., 01:35:45 or 11:26). Seconds will only be reported if they were measured as part of the study.
- The last footnote will be
  - "PROGRAM SOURCE: ...\\xx.sas, DATA CUT OFF DATE: DDMMMYYYY, RUN DATE: DDMMYY hh:mm".

  where the extract date is the datestamp of the data snapshot used.

15.2 Appendix B: BRISTOL STOOL CHART



15.3 Appendix C: Table, Listing and Figure Specifications

Specifications for tables, listings and figures are included in a separate document that may be updated independently and tracked as needed until database lock. The file approved with this SAP is v2.0 dated 16SEP2020.