Statistical Analysis Plan: C3718-202

Version 2.0, 15 Jun 2017

Study Title:	A Phase 2b, Randomized, Double-blind, Placebo-controlled, Parallel-group, Dose-range-finding Trial of IW-3718 Administered Orally for 8 Weeks to Patients with Symptomatic Gastroesophageal Reflux Disease Not Completely Responsive to Proton Pump Inhibitors
Study Number:	ICP-3718-202
Study Phase:	2b
Sponsor:	Ironwood Pharmaceuticals, Inc. 301 Binney Street Cambridge, MA 02142
Final Date:	15 Jun 2017

Confidentiality Statement

The contents of this document are confidential and belong to Ironwood Pharmaceuticals, Inc. Except as may be otherwise agreed to in writing, by accepting or reviewing these materials, you (including any colleagues or subordinates) agree to hold such information in confidence and not to disclose it to others (except where required by applicable law), nor to use it for unauthorized purposes. In the event of actual or suspected breach of this obligation, Ironwood should be promptly notified.

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Full Term
AE	adverse event
ANCOVA	analysis of covariance
ANOVA	analysis of variance
BID	twice daily
BMI	body mass index
BP	blood pressure
BPM	beats per minute
CMH	Cochran-Mantel-Haenszel
DHSS	daily heartburn severity score
DRFS	daily regurgitation frequency score
ECG	electrocardiogram
eCRF	electronic clinical report form
eDiary	electronic diary
EGD	esophagogastroduodenoscopy
EOT	end of treatment
EQ	EuroQol
GERD	gastroesophageal reflux disease
GSRS	the Gastrointestinal Symptoms Rating Scale
ICF	informed consent form
IPD	important protocol deviation
LLN	lower limit of normal
LOCF	last observation carried forward
LS	least-squares
OC	observed cases
MedDRA	the Medical Dictionary for Regulatory Activities
mITT	modified Intent-to-Treat
MMRM	mixed model for repeated measures
mRESQ-eD	the modified Reflux Symptom Questionnaire Electronic Diary
PCS	potentially clinically significant

Abbreviation	Full Term
PID	patient identification
PPI	proton pump inhibitor
PP	per-protocol
PT	preferred term
QD	once daily
QIDS	the Quick Inventory of Depressive Symptomatology
ROC	receiver operating characteristic
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SI	Le Système International d'Unités (International System of Units)
SOC	system organ class
TEAE	treatment-emergent adverse event
ULN	upper limit of normal
WHO DDE	World Health Organization Drug Dictionary Enhanced
WHSS	weekly heartburn severity score
WRFS	weekly regurgitation frequency score

1. INTRODUCTION

This statistical analysis plan (SAP) provides a more technical and detailed elaboration of the statistical analyses of the efficacy and safety data as outlined and/or specified in the amended protocol of Study ICP-3718-202 (dated 27 July 2016). Specifications for the tables, figures, and data listings are contained in a separate document.

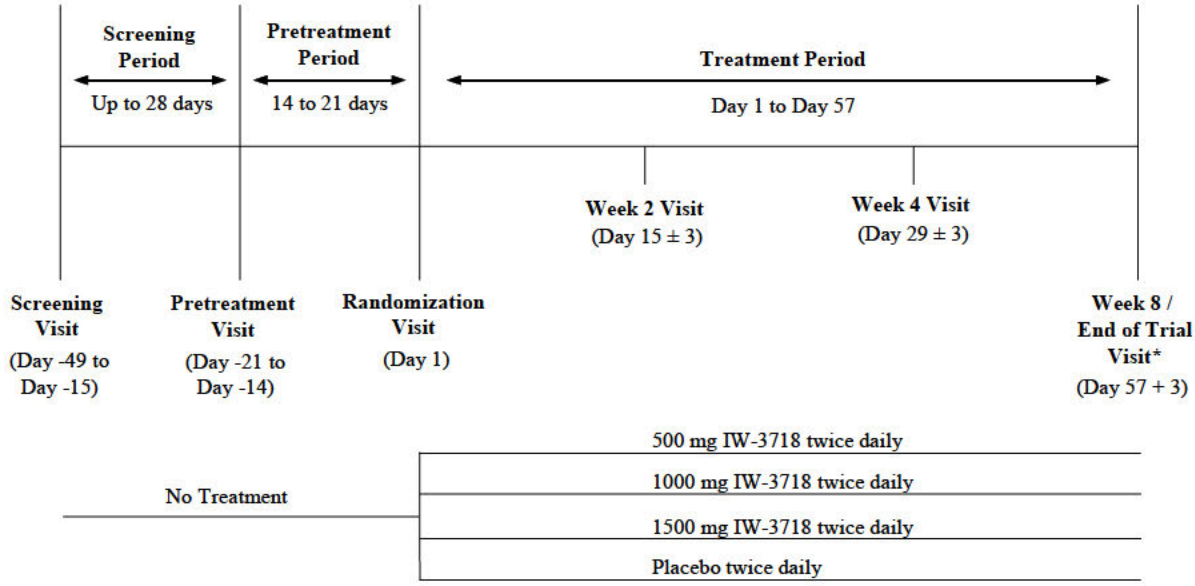
2. STUDY OBJECTIVES

The objectives of this study are to evaluate the safety, efficacy, and dose-response relationship of IW-3718 administered orally to patients who have gastroesophageal reflux disease (GERD) and continue to experience GERD symptoms while receiving once-daily (QD), optimized, standard-dose proton pump inhibitors (PPIs).

3. STUDY DESIGN

ICP-3718-202 is a multicenter, randomized, double-blind, placebo-controlled, parallel-group, 8-week study, consisting of 3 distinct periods as illustrated in Figure 3.0-1 below. The study will enroll patients who have GERD and continue to experience GERD symptoms while receiving QD, standard-dose PPI therapy that in the investigator's opinion has been optimized. Eligible patients will continue to take their PPI and will be randomized to placebo, 500 mg IW-3718 twice daily (BID), 1000 mg IW-3718 BID, or 1500 mg IW-3718 BID.

Figure 3.0-1. Overview of Study Design



Note: There is no Day 0

* This visit represents the end of the study

Screening Period: The Screening Period starts with the signature of the informed consent form (ICF) and may last for up to 28 days. During this period, patient eligibility for entry into the Pretreatment Period will be determined.

Pretreatment Period: The Pretreatment Period is defined as the 14 to 21 calendar days immediately before the Randomization Visit. During this period, patients will be required to use a handheld electronic diary (eDiary) to complete daily assessments (GERD symptoms, dyspepsia symptoms, assessment of sleep, and use of per-protocol rescue medicine) for at least 5 days each week during the last 14 calendar days before the Randomization Visit and weekly assessments

(symptom bothersomeness and degree of relief questions) at least once during the last 7 calendar days before the Randomization Visit in order to be eligible for randomization.

Treatment Period: The Treatment Period begins with treatment assignment and lasts for 8 weeks. Patients will be stratified by whether they have, or do not have, erosive esophagitis on the screening esophagogastroduodenoscopy (EGD) and randomly assigned to 1 of 4 treatments (1:1:1:1) within each stratum: placebo, 500 mg IW-3718 BID, 1000 mg IW-3718 BID, or 1500 mg IW-3718 BID. During this period, patients will continue to use the eDiary to provide their daily assessments (GERD symptoms, dyspepsia symptoms, assessment of sleep, and use of per-protocol rescue medicine) and weekly assessments (symptom bothersomeness, degree of relief, and treatment satisfaction questions).

4. STUDY ASSESSMENTS

4.1 EFFICACY ASSESSMENTS

The daily patient assessments used to determine the primary and secondary efficacy parameters are the daily assessment of heartburn symptoms (assessed on a 0-to-5 ordinal severity scale) obtained from the modified Reflux Symptom Questionnaire Electronic Diary (mRESQ-eD). Additional assessments will also be used to determine the other efficacy parameters, as described in the sections that follow.

4.1.1 Symptom Severity and Relief Assessments

4.1.1.1 Daily Assessments

During the Pretreatment and Treatment Periods, patients will enter information into their eDiary at approximately the same time each day.

- GERD symptoms (mRESQ-eD) completed once daily before going to bed each night.

  Note: The following items are assessed on a 0-5 severity scale: 0=Did not have, 1=Very mild, 2=Mild, 3=Moderate, 4=Moderately severe, 5=Severe

  - Heartburn
  - Burning feeling behind breastbone or in the center of the upper stomach
  - Pain behind the breastbone or in the center of the upper stomach
  - Difficulty swallowing
  - Hoarseness
  - Cough

  Note: The following items are assessed on a 0-4 frequency scale: 0=Never, 1= Rarely, 2=Sometimes, 3=Often, 4=Very often

  - Regurgitation (liquid or food) moving upwards toward your throat or mouth
  - An acid or bitter taste in the mouth
  - Burping
  - Coughing

- Dyspepsia symptoms completed once daily before going to bed each night.

  Note: All items are assessed on a 0-to-10 numerical rating scale (NRS) with 0=not having the symptom and 10=having the worst possible level of the symptom

- Worst nausea (feeling like you might throw up)
- Worst stomach fullness after you finished eating
- Difficulty you had finishing your meals because you felt full too quickly
- Worst abdominal pain

- Assessment of sleep completed once daily upon getting up each morning (5:00 a.m. to 12:00 p.m.)

  - Last night, did you wake up during the night after falling asleep? Y/N
  - [If yes], how many times did you wake up last night after falling asleep?
  - How long did you sleep last night? Do not count any time you lay in bed, but did not sleep.
  - Please rate the overall quality of your sleep last night. 1=very poor, 2=poor, 3=fair, 4=good, 5=very good

- Use of per-protocol rescue medicine (antacids)
- PPI administration

4.1.1.2 Weekly Assessments

The following information will be entered into the eDiary each week at about the same time as the evening daily assessment:

- Degree of relief questions

  Note: all items are assessed on a 7-point balanced ordinal scale: 1=Significantly relieved, 2=Moderately relieved, 3=Somewhat relieved, 4=Unchanged, 5=Somewhat worse, 6=Moderately worse, 7=Significantly worse

  - How would you rate your heartburn (a burning sensation in your chest, behind the breastbone) over the past 7 days?
  - How would you rate your regurgitation (the feeling of stomach contents, either liquid or food, moving upwards to your throat or mouth) over the past 7 days?
  - Compared to before you started this study, how would you rate your overall GERD symptoms over the past 7 days?

- Global Treatment Satisfaction Assessment

  Note: the following items will be assessed on a 5-point ordinal scale: 1=Very dissatisfied, 2=Dissatisfied, 3=Neither satisfied nor dissatisfied, 4=Satisfied, 5=Very satisfied

  - How would you rate your satisfaction with the study treatment?

- Bothersomeness Assessment

  Note: the following items will be assessed on a 5-point ordinal scale: 1=Not at all, 2=A little bit, 3=A moderate amount, 4=A great deal, 5=An extreme amount

  - How much were you bothered by heartburn (a burning sensation in your chest, behind the breastbone) over the past week?
  - How much were you bothered by regurgitation (the feeling of stomach contents, either liquid or food, moving upwards to your throat or mouth) over the past week?

4.1.2 GSRS-Self

The Gastrointestinal Symptoms Rating Scale (GSRS)-Self (see Appendix I) is a self-administered 15-item questionnaire [redacted]

[redacted]

4.1.3 End-of-Treatment Question

Patients will be asked a single item at their EOT visit, asking about the difficulty of swallowing the treatment medication. The item will ask "How difficult were the tablets to swallow?" and be rated on a 4-point ordinal scale:

1 = Not at all difficult
2 = A little difficult
3 = Moderately difficult
4 = Extremely difficult

4.1.4 QIDS-SR-16

The Quick Inventory of Depressive Symptomatology (QIDS)-SR-16 is a self-completed questionnaire [redacted]

4.1.5 mRESQ Cognitive Debriefing Interviews

A sample of patients at selected sites will be asked to participate in an optional symptom diary interview, either via phone or face to face. During the one-time interview, patients will be asked to respond to open-ended questions intended to assess the content validity of the mRESQ-eD instrument. The analysis of these cognitive debriefing interview data will be described in a separate document.

4.2 HEALTH OUTCOMES ASSESSMENTS

4.2.1 SF-12V2 Health Survey

The SF-12V2 (see Appendix II) is a widely used generic measure of health status [redacted]

4.2.2 EQ-5D-3L

The EuroQol (EQ)-5D-3L (see Appendix III) is a generic measure of health status widely used in Europe. [redacted]

[redacted]

4.3 SAFETY ASSESSMENTS

4.3.1 Adverse Events

An adverse event (AE) is any untoward medical occurrence in a patient. All AEs will be captured from the time the patient signs the informed consent until 7 days following the EOT Visit.

For all AEs, the investigator will provide an assessment of causal relationship to study medication (related or unrelated) and an assessment of the severity (mild, moderate, or severe). All AEs will also be categorized as serious or non-serious.

4.3.2 Medical History

A complete medical history will be performed as defined in the Schedule of Events (Table 4.4-1).

4.3.3 Physical Examination, Body Weight, and Height

A complete physical examination will be performed as defined in the Schedule of Events (Table 4.4-1). A physical examination will include the following assessments: general appearance; head, ears, eyes, nose, and throat; cardiovascular system; neck; respiratory system; musculoskeletal system; abdomen/liver/spleen; nervous system; lymph nodes; skin; neurologic status; and mental status. Breast, genitourinary, and rectal examinations are optional.

Each patient's weight will be recorded at every study visit; height will only be recorded at the Screening Visit.

4.3.4 Electrocardiograms

A 12-lead electrocardiogram (ECG) will be performed as defined in the Schedule of Events (Table 4.4-1).

4.3.5 Vital Signs

Vital sign measurements will be performed as defined in the Schedule of Events (Table 4.4-1). Vital sign measurements include oral temperature (°C), respiratory rate, systolic and diastolic blood pressure (BP), and pulse. Respiratory rate, pulse, and BP readings will be taken after the patient has been seated for at least 5 minutes.

4.3.6 Clinical Laboratory Determinations

Blood and urine samples for clinical laboratory tests will be collected at the visits defined in the Schedule of Events (Table 4.4-1). The clinical laboratory evaluations will include the clinical chemistry, hematology, coagulation, and urinalysis parameters presented in Table 4.3-1.

Table 4.3-1. Clinical Laboratory Tests

Clinical Chemistry	Hematology (CBC)	Complete Urinalysis
A1C	Hematocrit	pH and specific gravity
Albumin	Hemoglobin	Bilirubin
Alkaline phosphatase	Platelet count	Glucose
ALT	MPV	Ketones
AST	RBC count	Leukocytes
Bicarbonate	WBC count	Nitrites
BUN	WBC differential	Occult blood
Calcium	(% and absolute)	Protein
Chloride	Basophils	Urobilinogen
Total cholesterol	Eosinophils	
HDL cholesterol	Lymphocytes	
LDL cholesterol	Monocytes	
Creatinine	Neutrophils	
GGT	RBC indices	
Glucose	MCH	
Iron	MCHC	
LDH	MCV	
Magnesium	RDW	
Phosphorus	Coagulation	
Potassium	aPTT	
Sodium	PT	
Total bilirubin		
Total protein		
Triglycerides		
Uric acid		

Abbreviations: A1C = glycated hemoglobin; ALT = alanine aminotransferase; aPTT = activated partial thromboplastin time; AST = aspartate aminotransferase; BUN = blood urea nitrogen; CBC = complete blood count; GGT = gamma glutamyl transferase; HDL = high-density lipoprotein; LDH = lactate dehydrogenase; LDL = low-density lipoprotein; MCH = mean corpuscular hemoglobin; MCHC = mean corpuscular hemoglobin concentration; MCV = mean corpuscular volume; MPV = mean platelet volume; PT = prothrombin time; RBC = red blood cell; RDW = red cell distribution width; WBC = white blood cell

A blood sample for serum pregnancy testing will be collected from all female patients of childbearing potential at the Screening and EOT Visits, and a urine sample will be collected at the Randomization Visit (prior to dosing) and the Week 4 Visit. These pregnancy test results must be negative for patient eligibility. Ironwood may perform non-genetic analyses on existing plasma or serum samples for research purposes (e.g., bile in serum).

A urine screen for selected drugs of abuse (cocaine, barbiturates, amphetamines, opiates, benzodiazepines, and cannabinoids) and a serum alcohol screen will be performed at the Screening Visit.

4.3.7 Esophagogastroduodenoscopy

At all sites, an EGD will be performed on all patients at the screening visit. An additional EGD will be performed at the Week 8/EOT Visit in all patients who had an EGD during the Screening Period that demonstrated Grade C or D esophagitis.

4.3.8 Bravo

At all sites, all patients will undergo 48 to 96 hours of pH testing with the Bravo device. If for some reason 96 hours of testing is not possible, then approximately 48 hours of testing is acceptable. In addition, for the entire 48-96-hour Bravo pH monitoring period, patients will also complete a paper diary, recording at a minimum all instances of meals, snacks, drinks, and/or resting in the supine position.

4.3.9 Bilitec

At certain participating sites, selected patients who are screened with EGD and Bravo testing will also receive insertion of a Bilitec device during the same procedure. Patient selection will be dependent on Investigator judgement and patient consent. All such patients will return to the site approximately 24 hours after the procedure for removal of the device.

4.4 SCHEDULE OF EVENTS

The schedule of study procedures and assessments is presented by visit in the Schedule of Events in Table 4.4-1.

Table 4.4-1. Schedule of Events

	Screening Period (Up to 4 weeks)[w]	**Pretreatment Period (2 weeks)**	**Treatment Period (8 weeks)**				**Follow-up**
Visit Days →	Screening Visit (Day -49 to Day -15)	Pretreatment Visit (Day -21 to Day -1)	Randomization Visit (Day 1)	Week 2 Visit (Day 15 ± 3)	Week 4 Visit (Day 29 ± 3)	Week 8 / End-of-treatment Visit (Day 57 + 3)	EOT + 7 days
Visit Numbers →	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	
Study Procedure ↓							
Inclusion and Exclusion Criteria Verification	X	X	X				
Signing of ICF	X						
IWRS Registration (a)	X	X	X	X	X	X	
Demographics	X						
Medical History	X						
Physical Examination (b)	X					X	
Body Weight and Height (c)	X	X	X	X	X	X	
H2RA/Antacid Washout (d)	X						
Bilitec Monitoring (e)	X						
EGD (f)	X					X	
48 to 96 Hours of pH Testing with Bravo Device (g)	X						
Seated Vital Signs (h)	X	X	X	X	X	X	
12-Lead ECG (i)	X		X			X	
Prior and Concomitant Medications (j)	X	X	X	X	X	X	

Table 4.4-1. Schedule of Events

	Screening Period (Up to 4 weeks)[w]	**Pretreatment Period (2 weeks)**	**Treatment Period (8 weeks)**				**Follow-up**
Visit Days →	Screening Visit (Day -49 to Day -15)	Pretreatment Visit (Day -21 to Day -1)	Randomization Visit (Day 1)	Week 2 Visit (Day 15 ± 3)	Week 4 Visit (Day 29 ± 3)	Week 8 / End-of-treatment Visit (Day 57 + 3)	EOT + 7 days
Visit Numbers →	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	
Study Procedure ↓							
Clinical Laboratory Tests (k)	X		X		X	X	
Serum Pregnancy Test (l)	X					X	
Urine Pregnancy Test (l)			X		X		
Drug and Alcohol Screen (m)	X						
AE Evaluations (n)	X	X	X	X	X	X	
Rescue Medicine Dispensed (o)		X	X	X	X		
PDA Training and Dispensation		X					
Saliva Collection (p)	X	X	X		X	X	
eDiary (q)		X	X	X	X	X	
Weekly Symptom and Treatment Assessments (r)		X	X	X	X	X	
GSRS-Self			X		X	X	
QIDS-SR-16			X				
SF-12V2			X		X	X	
EQ-5D-3L			X		X	X	
Randomization			X				

Table 4.4-1. Schedule of Events

	Screening Period (Up to 4 weeks)[w]	**Pretreatment Period (2 weeks)**	**Treatment Period (8 weeks)**				**Follow-up**
Visit Days →	Screening Visit (Day -49 to Day -15)	Pretreatment Visit (Day -21 to Day -1)	Randomization Visit (Day 1)	Week 2 Visit (Day 15 ± 3)	Week 4 Visit (Day 29 ± 3)	Week 8 / End-of-treatment Visit (Day 57 + 3)	EOT + 7 days
Visit Numbers →	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	
Study Procedure ↓							
Study Medication Dispensed (s)			X	X	X		
Study Medication Return (t)				X	X	X	
PDA Return						X	
mRESQ-eD Debrief Interview (x)		X (x)			X (x)		
End of Treatment Question (u)						X	
Safety Follow-up Call (v)							X

Abbreviations: AE = adverse event; BP = blood pressure; CBC = complete blood count; ECG = electrocardiogram; EGD = esophagogastroduodenoscopy; EOT = End-of-treatment; EQ = EuroQol; GSRS = Gastrointestinal Symptoms Rating Scale; H2RA = histamine-2 receptor antagonist; HEENT = head, eyes, ears, nose, and throat; ICF = informed consent form; IWRS = interactive web response system; PDA = personal digital assistant; PPI = proton pump inhibitor; QID = Quick Inventory of Depressive Symptomology; mRESQ-eD = modified Reflux Symptom Questionnaire Electronic Diary; SAE = serious adverse event; SF = short form

a. Site personnel will interact with IWRS to register the patient visit and transition the patient to the next appropriate study period.
b. A physical examination should include the following assessments: general appearance; HEENT; neck; cardiovascular; respiratory; abdomen/liver/spleen; musculoskeletal; lymph nodes; skin; neurologic; nervous system, and mental status. Breast, genitourinary, and rectal examinations are optional and may be performed at the discretion of the Investigator.
c. Height will be measured only at the Screening Visit.
d. During the Screening Period, patients will begin to washout H2RAs and antacids. H2RAs should be stopped 5 calendar days prior to the EGD and Bravo pH monitoring and antacids should be stopped 1 calendar day prior to the EGD and Bravo pH monitoring. Patients may resume antacid use upon completion of the Bravo testing, but must refrain from H2RA use for the remainder of the study. During the Pretreatment Period, patients will refrain from using any anti-reflux medications, antacids, and H2RAs, except for the antacids that are provided as rescue medicine). Patients will continue to use their current PPI during the Pretreatment Period.

e. At selected sites, selected patients who are screened for the study with EGD and Bravo testing will also be given the option of having a Bilitec monitor inserted during the same procedure. These patients will return 24 hours later for removal of the probe. The results of the Bilitec test will not affect qualification for enrollment. The Bilitec device and Bravo recorder internal clocks must be synchronized, and both devices should be activated concurrently; however, if it is not possible to simultaneously activate both devices, the Bravo recorder should be activated first, with the Bilitec device activation immediately (within 5 minutes) after the Bravo pH recorder activation.
f. All patients will be required to undergo an EGD during the Screening Period. There must be a minimum of 7 days between the EGD and the start of the Pretreatment Period to allow for pH testing and patient stabilization. An EGD will be performed at the Week 8 / EOT Visit in all patients who have Grade C or D esophagitis (based on the Los Angeles classification of esophagitis) on the EGD obtained during the Screening Period.
g. Approximately 48 to 96 hours of pH testing with the Bravo device. If for some reason 96 hours of testing is not possible, then approximately 48 hours of testing is acceptable. During the entire period of Bravo pH monitoring (approximately 48-96 hours), patients will record ingestion of anything other than water by depressing the 'meal' button on the Bravo recorder upon the beginning and completion of the meal, snack, or drink. Patients will also record any periods during which they are lying down by depressing the 'supine' button on the Bravo recorder at the beginning and upon completion of the supine period. In addition, for the entire 48-96-hour Bravo pH monitoring period, patients will complete a paper diary, recording at a minimum all instances of meals, snacks, drinks, and/or resting in the supine position.
h. Vital sign measurements include oral temperature (°C), respiratory rate, systolic and diastolic BP, and pulse. Respiratory rate, BP, and pulse measurements must be obtained after the patient has been seated for at least 5 minutes.
i. 12-Lead ECGs should be obtained after the patient has been supine for at least 5 minutes.
j. Prior medications will be collected at the Screening Visit as follows: all medicines taken by the patient during the 30 days before the Screening Visit, most recent use of an H2RA, and most recent use of an antacid.
k. Clinical laboratory tests include clinical chemistry, hematology (CBC), coagulation, and urinalysis. If the triglyceride value exceeds the protocol-specified criteria and the patient was not under fasted conditions, the patient may return to complete a fasted lipid panel.
l. For all female patients of childbearing potential, a negative serum pregnancy test must be documented at the Screening Visit, and a negative urine pregnancy test must be documented at the Randomization Visit (before dosing) in order for the patient to be randomized into the study. A urine pregnancy test will be obtained at the Week 4 Visit; a serum pregnancy test will be conducted at the EOT Visit.
m. Patients must undergo a urine drug screen for selected drugs of abuse (cocaine, barbiturates, amphetamines, opiates, benzodiazepines, and cannabinoids) and a serum alcohol screen at the Screening Visit.
n. All AEs will be captured from the time the patient signs the ICF through the EOT Visit.
o. Rescue medicine will be supplied to patients at the Pretreatment Visit, and if needed, at subsequent visits.
p. At each of the indicated visits, patients will provide approximately 1 mL of saliva for future use (bile acids will be quantified in the saliva, and the quantity of bile acid may be used to define potential responders to IW-3718). At the time of saliva collection, study site staff will collect the following information: time of day, time of patient's last meal, and time of last study medication administration. During the Randomization Visit, study site staff will collect the saliva sample immediately following the light snack but prior to study drug administration.
q. The eDiary will be dispensed at the Pretreatment Visit and patients must complete at least 5 days each week during the 14 days before the Treatment Period in order to be eligible for randomization. Patients should bring their eDiary to each visit. The eDiary will collect daily PPI administration, rescue medication use, mRESQ-eD (daily), Daily Assessment of Sleep (daily), Daily Dyspepsia Symptoms (daily).
r. Symptom bothersomeness and symptom relief items (weekly); Treatment satisfaction (weekly).
s. The first dose of study medication will be administered in the clinic with liquid and a snack at the Randomization Visit. At all other visits, patients will take their study medication prior to arriving at the clinic, but will be dispensed additional doses needed until the next study visit.
t. Treatment compliance with study drug will be assessed based on return of unused tablets.
u. All patients will be asked an EOT question regarding the difficulty of swallowing the IW-3718 or placebo tablets.

v. Study site will contact each patient via telephone 7 days after the EOT Visit to collect information pertaining to ongoing AEs/SAEs and information concerning any new AEs/SAEs since the EOT Visit.
w. Ironwood may grant a one-week extension of the screening period window if needed for logistical delays (e.g., subject travel, scheduling issues, delays in test results, equipment malfunction, etc.). Approval should be requested from Ironwood prior to each extension.
x. Patients participating in the optional mRESQ-eD Cognitive Debriefing interviews will be assigned to one of three interview groups. Interviews will take place at different timepoints depending on the patient's group assignment.

5. STATISTICAL METHODS

5.1 GENERAL CONSIDERATIONS

All statistical analyses will be performed using SAS® Version 9.3 (or later) for Windows.

5.1.1 Data Display and Summary

All summaries will be presented by treatment unless specified otherwise. Descriptive statistics including the number of patients, mean, standard deviation (SD), median, minimum, and maximum will be calculated for continuous variables. Frequencies and percentages for each category will be calculated for categorical variables. Percentages will be based on the total number of non-missing values. The number missing will be presented, but without a percentage.

5.1.2 Handling of Missing Values

Unless stated otherwise, missing data will not be imputed.

5.1.3 Statistical Significance and Multiplicity

All hypothesis tests will be two-sided with a 5% significance level, and 95% confidence intervals will be used, unless stated otherwise. No adjustments for multiplicity are planned; p-values will be reported as nominal.

5.1.4 Interim Analysis and Data Monitoring

No interim analyses are planned.

5.1.5 Multicenter Studies

This study is being conducted in approximately 70 centers. Due to the potential of small numbers of patients per center, data will be pooled by center into the following 5 geographic regions (as listed in Table 5.1-1): Northeast, Southeast, Midwest, Southwest, and West. All analyses using trial center will use this 5-category geographic region variable.

Table 5.1-1. Definition of Geographic Regions

Northeast	*Southeast*	*Midwest*	*Southwest*	*West*
CT	AL	IA	AZ	CA
DE	AR	IL	NM	CO
MA	FL	IN	OK	ID
MD	GA	KS	TX	MT
ME	KY	MI		NV
NH	LA	MN		OR
NJ	MS	MO		UT
NY	NC	ND		WA
PA	SC	NE		WY
RI	TN	OH		
VT	VA	SD		
	WV	WI		

5.2 ANALYSIS POPULATIONS

5.2.1 Screened Population

The Screened Population consists of all patients who signed informed consent and received a patient identification (PID) number.

5.2.2 Randomized Population

The Randomized Population consists of all patients in the Screened Population who were randomized to a treatment group in the study.

5.2.3 Modified Intent-to-Treat Population

The modified Intent-to-Treat (mITT) Population consists of all patients in the Randomized Population who received at least one dose of study treatment.

5.2.4 Per-Protocol Population

The Per-Protocol (PP) Population consists of all patients in the mITT Population who have a minimum of 6 weeks of eDiary data for the heartburn severity and regurgitation frequency scores and > 80% compliance with study treatment for the available Treatment Period days.

5.2.5 Safety Population

The Safety Population consists of all patients in the Randomized Population who received at least one dose of study treatment.

5.3 PROTOCOL DEVIATIONS

Protocol deviations and Significant Protocol Deviations (SPD) will be identified and documented for all randomized patients prior to unblinding. The number and percentage of subjects with SPDs will be presented by study treatment and SPD category. Listings of protocol deviations and SPDs will also be provided.

5.4 PATIENT DISPOSITION

The number and percentage of patients included in the mITT, PP, and Safety Populations will be presented overall and by treatment group for the Randomized Population. The number of patients in the Screened Population will be presented overall.

The number and percentage of screen failures (i.e., patients who entered the Screening Period but not the Pretreatment Period) and patients ineligible for randomization (i.e., patients who entered the Pretreatment Period but were not randomized), along with the associated reasons for failure or ineligibility, will be tabulated overall for the Screened Population.

The number and percentage of patients who completed treatment, who completed the study, and who prematurely discontinued (overall and by reason for discontinuation) will be presented for each treatment group and overall for the Randomized Population.

5.5 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographic parameters (age; age group; race; ethnicity; sex; weight; height; and body mass index [BMI], calculated as weight [kg]/(height [m])^2) will be summarized by treatment group for the Randomized, mITT, and PP Populations. Baseline efficacy assessments (including GERD symptoms, dyspepsia symptoms, assessments of sleep, bothersomeness assessments, and the degree of relief assessments) will be summarized by treatment group for the mITT and PP Populations.

Abnormalities in patients' medical and surgical histories will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), version 18.1. The number and percentage of patients with abnormalities in medical and surgical histories in each system organ class (SOC) and preferred term (PT) will be summarized by treatment group for the Safety Population.

5.6 MEASUREMENT OF TREATMENT COMPLIANCE

The first dose of study medication will be taken in clinic at the Randomization Visit (Day 1). Patients should take their second dose that evening immediately upon completion of dinner. During the Treatment Period, patients will take study medication BID at home, in the morning (immediately upon completion of breakfast) and in the evening (immediately upon completion of dinner), even on study visit days. The last dose of study medication will be taken the morning of the EOT Visit. The double-dummy dosing requires all patients to take three tablets for each dose.

Dosing compliance for the Treatment Period is defined as the number of tablets taken by a patient during the Treatment Period divided by the number of tablets that were expected to be taken during the Treatment Period multiplied by 100. The total number of tablets taken will be calculated as total number of tablets dispensed minus total number of tablets returned minus total number of tablets lost. The total number of tablets expected to be taken during the Treatment Period is the number of days between the date of first dose taken and the date of last dose taken, inclusive, multiplied by 6 (multiplied by 6 minus 3 if the last dose is a morning dose). Patients will be included in the period summary only for the periods they were on study (i.e., the periods that are prior to or contain the last dose date).

Descriptive statistics for study drug compliance will be presented for the Safety Population by treatment group.

Patients will record their PPI administration each day (once daily) in their eDiary. PPI compliance for a specified period is defined as the number of days for which a patient answered Yes to the PPI administration question in the eDiary divided by the total number of days for which the patient provided an answer (Yes or No) to the PPI administration question in that period multiplied by 100. PPI compliance will be summarized by treatment group for the Safety Population.

5.7 EXTENT OF EXPOSURE

Exposure to study drug for the Safety Population during the double-blind treatment period will be summarized by treatment group in terms of treatment duration, calculated as the number of days from the date of the first dose of double-blind study drug to the date of the last dose of double-blind study drug, inclusive. Treatment duration will also be categorized as <7 days; ≥7 days and <14 days; ≥14 days and <28 days; ≥28 days and <42 days; ≥42 days and <56 days; ≥56 days.

Patient-years, defined as exposure to the study drug in years, will be summarized by treatment group for the Safety Population.

5.8 PRIOR AND CONCOMITANT MEDICATION

Prior medication is defined as any medication taken before the date of the first dose of double-blind study drug. Concomitant medication is defined as any medication taken on or after the date of the first dose of double-blind study drug during the Treatment Period. Medications started after the date of the last dose of double-blind study drug will not be considered concomitant medications.

Both prior and concomitant medications will be coded by drug name and therapeutic class using World Health Organization Drug Dictionary Enhanced (WHO DDE) December 2015. The use of prior and concomitant medications will be summarized for the Safety Population by the number and percentage of patients in each treatment group receiving each medication within each therapeutic class. If a patient took a specific medication multiple times or took multiple medications in the same category (based on Anatomical-Therapeutic-Chemical classification), that patient would be counted only once for the coded drug name or therapeutic class.

5.9 EDIARY COMPLIANCE

Patients are required to enter their diary assessments into the eDiary before going to bed each night and upon getting up each morning during the Pretreatment and Treatment Periods. A daily report is considered complete if all questions in both the morning diary and evening diary (excluding the weekly assessments) are answered, with the following exceptions.

- For the first day of the Pretreatment Period, only the evening diary is expected.
- For the last day of the Pretreatment Period, only the morning diary is pertinent.
- For the first day of the Treatment Period, only the evening diary is pertinent.
- For the EOT Visit day, only the morning diary is expected.

eDiary compliance will be calculated for baseline (the Pretreatment Period), the Treatment Period, and weekly. For each period, eDiary compliance will be calculated as the number of completed daily reports divided by the expected number of daily reports in the period multiplied by 100. eDiary compliance will be summarized by treatment group for the mITT Population. In addition, the number and percentage of patients with ≥80% vs. <80% eDiary compliance for each period, as well as the number and percentage of patients with ≥4 vs. <4 complete daily reports each week, will be presented by treatment group for the mITT Population.

5.10 EFFICACY ANALYSIS

Unless otherwise specified, efficacy analyses will be based on both the mITT and the PP Populations.

Baseline values for efficacy parameters will be derived from the eDiary and/or electronic clinical report form (eCRF) data collected for the last week of the Pretreatment Period (or Week -1), specifically the period from 7 days prior to the day of randomization up to the time of randomization. If data for Week -1 is not available, data for Week -2 (the second last week of the Pretreatment Period) will be used.

For continuous efficacy parameters, the analysis value for an analysis period is defined as the average of the non-missing values during that period.

For weekly responder parameters based on daily eDiary assessments, a patient who missed 4 or more daily assessments during an analysis week will not be considered a responder for that week.

Table 5.10-1 provides the analysis time windows allowed for efficacy analyses based on daily assessments and Table 5.10-2 provides the analysis time windows allowed for efficacy analyses based on weekly assessments.

Table 5.10-1. Analysis Time Windows for Efficacy Analysis – Daily Assessments

Period	Analysis Week	Begins[a]	Ends[a]
Pretreatment (Baseline[b])	Week -2	Day -14	Day -8
	Week -1	Day -7	Day -1 (Day before randomization) [c]
Treatment	Week 1	Day 1 (Day of Randomization) [d]	Day 7
	Week 2	Day 8	Day 14
	Week 3	Day 15	Day 21
	Week 4	Day 22	Day 28
	Week 5	Day 29	Day 35
	Week 6	Day 36	Day 42
	Week 7	Day 43	Day 49
	Week 8	Day 50	Day 56
	Week 9	Day 57	Day 63
	•••	•••	•••

a. Relative to the date of randomization (Day 1).
b. Baseline values for efficacy parameters will be derived from the daily eDiary collected in the Pretreatment Period Week -1 (Week -2 if data from Week -1 not available)
c. Day 1 for efficacy parameters derived from the morning diary.
d. Day 2 for efficacy parameters derived from the morning diary.

Table 5.10-2. Analysis Time Windows for Efficacy Analysis – Weekly Assessments

Period	Analysis Week	Begins[a]	Ends[a]
Pretreatment (Baseline[b])	Week -2	Day -11	Day -5
	Week -1	Day -4	Day -1 (Day before randomization)
Treatment	Week 1	Day 7	Day 10
	Week 2	Day 11	Day 17
	Week 3	Day 18	Day 24
	Week 4	Day 25	Day 31
	Week 5	Day 32	Day 38
	Week 6	Day 39	Day 45
	Week 7	Day 46	Day 52
	Week 8	Day 53	Day 59
	Week 9	Day 60	Day 66
	•••	•••	•••

a. Relative to the date of randomization (Day 1).
b. Baseline values for efficacy parameters will be derived from the eDiary collected in the Pretreatment Period Week -1 (Week -2 if data from Week -1 not available)

5.10.1 Primary Efficacy Parameter

The primary efficacy parameter is the percent change from baseline (i.e., the Pretreatment Period) to Week 8 in weekly heartburn severity score (WHSS).

The WHSS for an analysis week is defined as the average of available daily heartburn severity scores (DHSS) during that week. DHSS is defined as the maximum of the 3 items measuring heartburn (Heartburn, Burning feeling behind chestbone or in the center of the upper stomach, and Pain behind chestbone or in the center of the upper stomach) from a particular day collected with the mRESQ-eD instrument. For a handful of patients who were assigned to an older version of the instrument, Burning feeling behind the chestbone and Burning feeling in the center of the upper stomach were assessed separately. For analysis purposes, the maximum of the two will be used to represent the item of Burning feeling behind chestbone or in the center of the upper stomach. Similarly, Pain behind chestbone and Pain in the center of the upper stomach were assessed separately for these patients. For analysis purposes, the maximum of the two will be used.

Percent change from baseline to Week 8 in WHSS (both last-observation-carried-forward [LOCF] and observed-case [OC]) will be summarized by treatment group. Each IW-3718 group will be compared with the placebo group using an analysis of covariance (ANCOVA) model with treatment group and esophagitis status (erosive esophagitis or no erosive esophagitis) as fixed-effect terms and baseline WHSS as a covariate. The least-squares (LS) mean for each treatment group, LSM difference between each IW-3718 group and the placebo group along with the corresponding confidence interval, and p-values for the pairwise comparisons vs. placebo will be presented. In addition, a linear contrast among the LS means for treatment groups will be conducted to test the overall ordinal dose response.

The treatment by esophagitis status interaction will be explored (using an ANCOVA model with an additional term for the interaction and/or plotting treatment group means by esophagitis status) to evaluate whether quantitative or qualitative interaction is present. If warranted, treatment comparisons within each esophagitis status will be conducted.

As a sensitivity analysis, percent change from baseline to Week 8 in WHSS will be evaluated employing a mixed model for repeated measures (MMRM) framework with week (categorical), treatment group, week-by-treatment group, and esophagitis status as fixed-effect terms and baseline WHSS as a covariate. A compound symmetry covariance structure will be used to model the covariance of within-patient results. Exploratory analyses with unstructured and autoregressive (1) covariance structures will also be considered.

The cumulative distribution function (CDF) of percent change from baseline to Week 8 in WHSS will be plotted by treatment group. To aid in the interpretation of the graphical presentation, a two-sample Kolmogorov-Smirnov test will be conducted between each IW-3718 group and the placebo group.

5.10.2 Secondary Efficacy Parameters

The secondary efficacy parameters include the following:

1. Percent change from baseline to Week 4 in WHSS
2. Change from baseline to Week 8 in WHSS
3. Change from baseline to Week 4 in WHSS

4. Proportion of patients who are overall heartburn responders

   An overall heartburn responder is a patient who is a weekly heartburn responder for at least 4 of the 8 treatment weeks and for at least 1 of the final 2 treatment weeks (i.e., Week 7 and Week 8). A weekly heartburn responder is a patient with a decrease of ≥ 30% from baseline in WHSS.
5. Proportion of patients with a DHSS of no more than very mild (≤ 1) on any day during Week 8
6. Proportion of patients with a DHSS of no more than very mild (≤ 1) on any day during Week 4
7. The number of days where DHSS was no more than very mild (≤ 1) during Week 8
8. The number of days where DHSS was no more than very mild (≤ 1) during Week 4
9. Change from baseline in the weekly average of each mRESQ-eD item by week.
10. Proportion of heartburn-free days during Week 8

    A heartburn free day is a day where DHSS=0.
11. Proportion of heartburn-free days during Week 4

The continuous parameters (#1, 2, 3, 7, 8, 9, 10, and 11) will be analyzed employing similar methods as described for the primary efficacy parameter.

For the analysis of responder parameters (#4, 5, and 6), the counts and proportion of responders will be calculated for each treatment group. The proportions of responders between each IW-3718 group and the placebo group will be compared using a Cochran-Mantel-Haenszel (CMH) test controlling for esophagitis status. The CMH test is the primary analysis for responder parameters. The difference in the proportion of responders between each IW-3718 group and the placebo group as well as the CMH estimates of odds ratio (IW-3718 over placebo) will be presented, along with the corresponding confidence intervals.

Corresponding to secondary efficacy parameters 7, 8, 10, and 11, additional analyses will be provided for the following parameters: 7a. Proportion of Days where DHSS was no more than very mild (≤ 1) during the 8-week Treatment Period; 8a. Proportion of Days where DHSS was no more than very mild (≤ 1) during the first 4 weeks of the Treatment Period; 10a. Proportion of heartburn-free days during the 8-week Treatment Period; 11a. Proportion of heartburn-free days during the first 4 weeks of the Treatment Period.

5.10.3 Other Efficacy Parameters

Other efficacy parameters include the following:

1. Percent change from baseline to Week 8 in weekly regurgitation frequency score (WRFS) in patients with a baseline WRFS of ≥ 2 (sometimes)

   The WRFS for an analysis week is defined as the average of available daily regurgitation frequency scores (DRFS) during that week. DRFS is defined as the maximum of the 2 items measuring regurgitation (Regurgitation [liquid or food moving upwards toward your throat or mouth] and An acid or bitter taste in the mouth) from a particular day collected with the mRESQ-eD instrument.
2. Percent change from baseline to Week 4 in WRFS in patients with a baseline WRFS of 2 (sometimes)
3. Change from baseline to Week 8 in WRFS
4. Change from baseline to Week 4 in WRFS
5. Proportion of patients who are overall regurgitation responders (of patients with a baseline WRFS of ≥ 2 [sometimes])

   An overall regurgitation responder is a patient who is a weekly regurgitation responder for at least 4 of the 8 treatment weeks and for at least 1 of the final 2 treatment weeks (i.e., Week 7 and Week 8). A weekly regurgitation responder is a patient with a decrease of ≥ 30% from baseline in WRFS.
6. Proportion of patients with a DRFS of no more than rarely (≤ 1) on any day during Week 8
7. Proportion of patients with a DRFS of no more than rarely (≤ 1) on any day during Week 4
8. Change from baseline to Week 8 in the number of days per week where DRFS was no more than rarely (≤ 1)
9. Change from baseline to Week 4 in the number of days per week where DRFS was no more than rarely (≤ 1)
10. Change from baseline to Week 8 in the weekly average of each of the daily symptom assessments
11. Weekly means for daily sleep assessments (number of awakenings, hours of sleep, sleep quality)
12. Weekly means for weekly symptom bothersomeness assessments, symptom relief assessments, and treatment satisfaction assessment

For analysis of continuous parameters (#1, 2, 3, 4, 8, 9, 10, 11, and 12), summary statistics will be presented by treatment group. Each IW-3718 group will be compared with the placebo group using an ANCOVA model with treatment group and esophagitis status as fixed-effect terms and

the corresponding baseline efficacy parameter value as a covariate. Note that for parameters without a baseline value (e.g., Treatment Satisfaction), an ANOVA model (in lieu of ANCOVA) will be used with treatment group and esophagitis status as fixed-effect terms.

For analysis of responder parameters (#5, 6, and 7), the counts and proportion of responders will be calculated for each treatment group. The proportions of responders between each IW-3718 group and the placebo group will be compared using a CMH test controlling for esophagitis status.

Corresponding to efficacy parameters 8 and 9, additional analyses will be provided for the following parameters: 8a. Change from baseline to the 8-week Treatment Period in proportion of days where DRFS was no more than rarely (≤ 1); 9a. Change from baseline to the first 4 weeks of Treatment Period in proportion of days where DRFS was no more than rarely (≤ 1).

Change/percent change from baseline in WHSS, WRFS, and the weekly average of each of the individual daily symptom assessment items (GERD & dyspepsia) will be summarized/analyzed by week, and the LS mean changes/LS mean percent changes will be plotted across treatment weeks.

The proportion of Degree of Relief responders (reported a score of 1-Significantly Relieved or 2-Moderately Relieved on the Degree of Relief question for at least 4 of the 8 treatment weeks) will be summarized by treatment group for heartburn, regurgitation, and overall GERD symptoms.

In addition, the GSRS-self will be summarized by treatment group by visit for each of the 15 items and for each of the 5 domains: abdominal pain (3 items), reflux symptoms (2 items), indigestion (4 items), diarrhea (3 items), and constipation (3 items). The EOT Visit difficulty of swallowing question will also be summarized by presenting the count and proportion of patients in each of the 4 categories (from Not at all Difficult to Extremely Difficult) by treatment group.

5.10.4 Exploratory Analyses

The following analyses intend to explore the relationship between measures of bile acid, acid reflux, and clinical response.

a. Sensitivity and specificity of bile acid levels measured in saliva as a test for bile acid pathophysiology with Bilitec test findings as the reference standard
b. Correlation between Bilitec test findings and Bravo pH monitoring results
c. Bilitec test findings as a predictor of response to treatment
d. Bile acid levels in saliva as a predictor of response to treatment
e. Bravo pH monitoring results as a predictor of response to treatment

Analyses (a) and (b) will be conducted based on the Screened Population; all other analyses will be conducted using the mITT Population.

5.10.5 Subgroup Analyses

Subgroup Analyses of efficacy parameters will be based on the mITT Population.

Esophagitis Status

Subgroup analysis by esophagitis status (erosive esophagitis vs. no erosive esophagitis at screening) will be performed for the following efficacy parameters.

- Percent change from baseline to Week 8 in WHSS
- Overall Heartburn Responder
- Percent change from baseline to Week 8 in WRFS
- Overall Regurgitation Responder
- Percent change from baseline to Week 8 in saliva bile acid level

Baseline WHSS

Subgroup analysis by baseline WHSS (≤3 vs. >3) will be performed for the following efficacy parameters.

- Percent change from baseline to Week 8 in WHSS
- Overall Heartburn Responder

Baseline WRFS

Subgroup analysis by baseline WRFS (≤3 vs. >3) will be performed for the following efficacy parameters.

- Percent change from baseline to Week 8 in WRFS
- Overall Regurgitation Responder

Baseline Bile Acid Level in Saliva

Subgroup analysis by baseline bile acid level in saliva (positive vs. negative) will be performed for the following efficacy parameters.

- Percent change from baseline to Week 8 in WHSS
- Overall Heartburn Responder
- Percent change from baseline to Week 8 in WRFS
- Overall Regurgitation Responder

Bilitec Monitoring Status

Subgroup analysis by Bilitec monitoring status (performed vs. not performed at screening) will be performed for the following efficacy parameters.

- Percent change from baseline to Week 8 in WHSS
- Overall Heartburn Responder
- Percent change from baseline to Week 8 in WRFS
- Overall Regurgitation Responder

Bilitec Bile Acid Status

Subgroup analysis by Bilitec bile acid status (positive vs. negative at screening) will be performed for the following efficacy parameters.

- Percent change from baseline to Week 8 in WHSS
- Percent change from baseline to Week 8 in WRFS

5.10.6 Rescue Medicine Use

Patients will report the frequency of their nighttime use of rescues medicine (liquid antacid) in the morning diary and their daytime use of rescue medicine in the evening diary. Nighttime use, daytime use, and overall daily use of rescue medicine will be summarized by treatment group for each treatment week and for the Treatment Period.

5.11 HEALTH OUTCOMES ANALYSIS

5.11.1 SF-12V2 Health Survey

Change from baseline to Week 4 and Week 8 in SF-12V2 (both LOCF and OC) will be summarized by treatment group for each of the 8 health concepts (physical functioning, role limitations due to physical health problems, bodily pain, general health, vitality [energy/fatigue], social functioning, role limitations due to emotional problems, and mental health [psychological distress and psychological well-being]) and for each of the two summary measures (the physical component and mental component).

5.11.2 EQ-5D-3L

Change from baseline to Week 4 and Week 8 in EQ-5D-3L (both LOCF and OC) will be summarized by treatment group for each of the 5 questions assessing the following domains: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Change from baseline to Week 4 and Week 8 in heath state as assessed by a 0 to 100 visual analogue scale (both LOCF and OC) will also be summarized by treatment group.

5.12 SAFETY ANALYSIS

Safety analyses will be performed using the Safety Population. Safety data summaries will be presented for each treatment group as well as for IW-3718 overall.

For analysis of safety parameters, the last non-missing value before the first dose of double-blind study drug will be used as baseline.

Table 5.12-1 below presents the visits assigned for the safety analysis corresponding to the range of trial days (window) during which an actual visit may have occurred.

Table 5.12-1. Visit Time Windows for Safety Analysis

Derived Visit	*Scheduled Test / Visit Day*[a]	*Window*
Baseline	Day 1	Days ≤ 1
Day 15 Visit	Day 15	Days [2, 22]
Day 29 Visit	Day 29	Days [23, 43]
Day 57 Visit	Day 57	Days ≥ 44
End of Treatment[b]	Final or termination visit	

a. Relative to the date of randomization; Day 1 = the day of randomization.
b. "End of Treatment" will be presented in analysis tables for safety parameters, including clinical laboratory and vital signs.

Test/Visit Day will be calculated as follows: test/visit date – date of randomization.

5.12.1 Adverse Events

Adverse events will be coded by system organ class and preferred term using the MedDRA dictionary, version 18.1.

An AE will be considered a treatment-emergent adverse event (TEAE) if it started after the first dose of study drug or started prior to the first dose of study drug but increased in severity after the first dose of study drug.

The incidence of TEAEs will be presented by system organ class and preferred term. The incidence of Severe TEAEs, TEAEs related to study drug, TEAEs leading to discontinuation of study drug, and treatment-emergent serious adverse events (SAEs) will also be presented, separately, by system organ class and preferred term. If a patient has more than 1 AE coded to the same preferred term, the patient will be counted only once for that preferred term using the highest severity and closest relationship to study drug.

The incidence of common (≥ 3% of patients in any treatment group) TEAEs will be presented using preferred terms by decreasing frequency in the IW-3718 overall group.

Data listings will be provided for deaths, SAEs, and AEs leading to discontinuation of study drug.

Additional analyses will be conducted on constipation as an AE of interest. Median time from study drug initiation to onset of the first occurrence of constipation will be estimated using the Kaplan-Meier method. Duration of treatment-emergent constipation will also be summarized. If a patient had multiple occurrences of treatment-emergent constipation, the occurrence with the longest duration will be used in the summary.

5.12.2 Clinical Laboratory Parameters

Clinical laboratory test values (in SI unit) will be summarized for each assessment time point. Change from baseline values will be summarized for each assessment time point post baseline.

Clinical laboratory test values will be considered potentially clinically significant (PCS) if they meet either the lower-limit or higher-limit PCS criterion listed in Table 5.12-2. The incidence of change in clinical laboratory values from non-PCS at baseline to PCS post baseline will be tabulated for patients with available non-PCS values at baseline and at least 1 assessment post baseline. A supportive listing of patients with PCS post-baseline values will be provided, including the PID number, trial center, baseline value, and all post-baseline (including non-PCS) values. A listing of all AEs for patients with post-baseline PCS laboratory values will also be provided.

Table 5.12-2. Criteria for Potentially Clinically Significant Laboratory Results

Parameter	*SI Unit*	*Lower Limit*	*Higher Limit*
CHEMISTRY			
Albumin	g/L	< 0.9 × LLN	> 1.1 × ULN
Alanine aminotransferase	U/L	—	≥ 3 × ULN
Alkaline phosphatase	U/L	—	≥ 3 × ULN
Aspartate aminotransferase	U/L	—	≥ 3 × ULN
Bicarbonate	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Bilirubin, total	µmol/L	—	> 1.5 × ULN
Calcium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Chloride	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Cholesterol, total	mmol/L	—	> 1.6 × ULN
Creatinine	µmol/L	—	> 1.3 × ULN
Glucose	mmol/L	< 0.8 × LLN	> 1.4 × ULN

Table 5.12-2. Criteria for Potentially Clinically Significant Laboratory Results

Parameter	*SI Unit*	*Lower Limit*	*Higher Limit*
Magnesium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Phosphate	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Potassium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Protein, total	g/L	< 0.9 × LLN	> 1.1 × ULN
Sodium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Urea nitrogen	mmol/L	—	> 1.2 × ULN
Uric acid	µmol/L	< 0.9 × LLN	> 1.1 × ULN
HEMATOLOGY			
Basophils, absolute cell count	10^9/L	—	> 3 × ULN
Eosinophils, absolute cell count	10^9/L	—	> 3 × ULN
Hematocrit	Ratio	< 0.9 × LLN	> 1.1 × ULN
Hemoglobin	g/L	< 0.9 × LLN	> 1.1 × ULN
Lymphocytes, absolute cell count	10^9/L	< 0.8 × LLN	> 1.5 × ULN
Mean corpuscular hemoglobin	pg	—	> 3 × ULN
Mean corpuscular hemoglobin concentration	g/L	—	> 3 × ULN
Mean corpuscular volume	fL	< 0.9 × LLN	> 1.1 × ULN
Monocytes, absolute cell count	10^9/L	—	> 3 × ULN
Neutrophils, absolute cell count	10^9/L	< 0.8 × LLN	> 1.5 × ULN
Platelet count	10^9/L	< 0.5 × LLN	> 1.5 × ULN
Red blood cell count	10^{12}/L	< 0.9 × LLN	> 1.1 × ULN
White blood cell count	10^9/L	< 0.7 × LLN	> 1.5 × ULN

LLN = lower limit of normal value provided by the laboratory; SI = *Le Système International d'Unités* (International System of Units); ULN = upper limit of normal value provided by the laboratory.

5.12.3 Vital Sign Parameters

Vital sign values will be summarized for each assessment time point. Change from baseline values will be summarized for each assessment time point post baseline.

Vital sign values will be considered PCS if they meet both the observed-value criterion and the change-from-baseline criterion listed in Table 5.12-3. The incidence of PCS vital signs will be tabulated for patients with a baseline value and at least 1 post-baseline assessment. A supportive listing of patients with PCS vital signs will be provided, including the PID number, trial center, baseline value, and all post-baseline (including non-PCS) values. In addition, a listing of all AEs that occurred in patients who had PCS vital sign values will be provided.

Table 5.12-3. Criteria for Potentially Clinically Significant Vital Signs

Parameter	***Flag***	***Criteria***[a]	
		Observed Value	*Change from Baseline*
Sitting systolic blood pressure, mm Hg	High	≥ 180	Increase of ≥ 20
	Low	≤ 90	Decrease of ≥ 20
Sitting diastolic blood pressure, mm Hg	High	≥ 105	Increase of ≥ 15
	Low	≤ 50	Decrease of ≥ 15
Sitting pulse rate, bpm	High	≥ 120	Increase of ≥ 15
	Low	≤ 50	Decrease of ≥ 15
Weight, kg	High	—	Increase of ≥ 7%
	Low	—	Decrease of ≥ 7%

a. A post-baseline value is considered potentially clinically significant if it meets both the observed-value and the change-from-baseline criteria.
bpm = beats per minute.

5.12.4 ECG Parameters

ECG values will be summarized for baseline and the EOT Visit. Change from baseline values will be summarized for the EOT Visit. ECG values will be categorized as low, normal, or high based on reference ranges provided by the site. Shifts from baseline to end of study will be tabulated for the following categories: normal, abnormal not clinically significant, and abnormal clinically significant. ECG values will be PCS if they meet the higher limit PCS criteria listed in Table 5.12-4. The incidence of change in ECG values from non-PCS at baseline to PCS post baseline will be tabulated for patients with non-PCS values at baseline and at least 1 assessment post baseline.

Table 5.12-4. Criteria for Potentially Clinically Significant ECG Parameters

ECG Parameter	Unit	Higher Limit
QRS duration	msec	≥ 150
PR interval	msec	≥ 250
QTc interval	msec	> 500

5.12.5 Other Safety Parameters

Any physical examination abnormality that the investigator considers to be potentially clinically significant will be reported as an AE. No separate analysis of Physical Examinations is planned.

5.13 DETERMINATION OF SAMPLE SIZE

The sample size per arm was determined by estimating the overall power of a linear trend test in a one-way design that included placebo and all the active treatment arms (500 mg BID, 1000 mg BID, and1500 mg BID of IW-3718). The efficacy endpoint of interest in the previous Phase 2a study (ICP-3718-201) was change from baseline in daytime heartburn defined as the presence of heartburn since the patient awoke that morning (assessed in the evening) and assessed with an 11-point [0 to 10] NRS, where 0=none and 10=very severe. In that study, IW-3718 demonstrated treatment differences near 0.70 points for change over the treatment period (e.g., -0.73 vs. -1.38 for placebo and IW-3718, respectively), and standard deviations near 1.35 for the LSMs. Employing these historical values, a study with 58 patients per arm (the expected number of patients at Week 8 given an enrollment of 260 patients) will have statistical power of at least 80% for the overall trend test (i.e., linear contrast) where the highest active treatment arm reflects 110% of the previously observed treatment difference and the lowest 2 active treatment arms reflect 55% of the same (two-sided, α=0.05). At the proposed sample size, a subsequent pairwise comparison between placebo and an active treatment arm reflecting the previously observed treatment difference will have 80% statistical power (two-sided, α=0.05).

6. CHANGES FROM ANALYSES PLANNED IN THE PROTOCOL

a. One additional analysis population (the Randomized Population) was defined in the SAP that was not previously defined in the Protocol.

b. Patient disposition will be summarized for the Randomized Population as opposed to the mITT Population as specified in the Protocol.

7. DATA HANDLING CONVENTIONS

7.1 REPEATED OR UNSCHEDULED ASSESSMENTS OF SAFETY PARAMETERS

If a patient has repeated assessments prior to the start of double-blind study drug, then the results from the final non-missing assessment made prior to the start of the double-blind study drug will be used as baseline. If end-of-study assessments are repeated or if unscheduled visits occur, the last non-missing post baseline assessment will be used as the end-of-study assessment for generating summary statistics. However, all post-baseline assessments will be used for PCS value determination as described above and all assessments will be presented in the data listings.

7.2 MISSING DATE OF THE LAST DOSE OF STUDY DRUG

When the date of the last dose of double-blind study drug is missing for a patient in the Safety Population, all efforts should be made to obtain the date from the Investigator. If after all efforts are made it is still missing, the last eDiary entry date will be used as the last dose date.

7.3 MISSING SEVERITY ASSESSMENT FOR ADVERSE EVENTS

If severity is missing for an AE that started prior to the first dose of double blind study drug, all efforts should be made to obtain the severity from the investigator. If it is still missing after all efforts, then a severity of "Mild" will be assigned. If the severity is missing for an AE that started after the first dose of double-blind study drug, then a severity of "Severe" will be assigned. The imputed values for the missing severity assessment will be used for the incidence summary, while the actual missing values will be presented in data listings.

7.4 MISSING RELATIONSHIP TO STUDY DRUG FOR ADVERSE EVENTS

If the relationship to study drug is missing for an AE that started after the first dose of double-blind study drug, all efforts should be made to obtain the relationship from the investigator. If it is still missing after all efforts, a causality of "Related" will be assigned in the corresponding analysis derived data set. The imputed values for the missing relationship to double-blind study drug will be used only for incidence summary, while the actual missing values will be presented in data listings.

7.5 MISSING DATE INFORMATION FOR ADVERSE EVENTS

The following imputation rules apply to cases in which the start date is incomplete (i.e., partial missing) for adverse events.

a. If AEs occurred during the Screening or Pretreatment Period (i.e., the Screening or Pretreatment Period is checked on the AE eCRF)

Missing day and month

- If the year is the same as the year of the date of informed consent, then the day and month of the date of informed consent will be assigned to the missing fields.
- If the year is prior to the year of the date of informed consent, then December 31 will be assigned to the missing fields.
- If the year is after the year of the date of informed consent, then January 1 will be assigned to the missing fields.

Missing month only

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

Missing day only

- If the month and year are the same as the month and year of the date of informed consent, then the date of informed consent will be assigned to the missing day.
- If either the year is before the year of the date of informed consent or if both years are the same but the month is before the month of the date of informed consent, then the last day of the month will be assigned to the missing day.
- If either the year is after the year of the date of informed consent or if both years are the same but the month is after the month of the date of informed consent, then the first day of the month will be assigned to the missing day.

If the stop date is complete and it is before the date of randomization and the imputed start date as above is after the stop date, the start date will be imputed by the stop date. If the stop date is complete and it is on or after the date of randomization and the imputed start date as above is after the stop date, the start date will be imputed by the date immediately before the date of randomization.

If the start date is completely missing, then the date of informed consent will be used to impute the start date.

b. If AEs occurred during the Treatment Period (i.e., the Treatment Period is checked on the AE eCRF)

Missing day and month

- If the year is the same as the year of the date of the first dose of double-blind study drug, then the day and month of the date of the first dose of double-blind study drug will be assigned to the missing fields.
- If the year is prior to the year of the date of the first dose of double-blind study drug, then December 31 will be assigned to the missing fields.
- If the year is after the year of the date of the first dose of double-blind study drug, then January 1 will be assigned to the missing fields.

Missing month only

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

Missing day only

- If the month and year are the same as the month and year of the date of the first dose of double-blind study drug, then the date of the first dose of double-blind study drug will be assigned to the missing day.
- If either the year is before the year of the date of the first dose of double-blind study drug or if both years are the same but the month is before the month of the date of the first dose of double-blind study drug, then the last day of the month will be assigned to the missing day.
- If either the year is after the year of the date of the first dose of double-blind study drug or if both years are the same but the month is after the month of the date of the first dose of double-blind study drug, then the first day of the month will be assigned to the missing day.

If the stop date is complete and the imputed start date as above is after the stop date, the start date will be imputed by the stop date.

If the start date is completely missing and the stop date is complete, then the following algorithm is used to impute the start date:

- If the stop date is after the date of the first dose of double-blind study drug, the date of the first dose of double-blind study drug will be assigned to the missing start date.
- If the stop date is before the date of the first dose of double-blind study drug, the stop date will be assigned to the missing start date.

7.6 MISSING DATE INFORMATION FOR PRIOR OR CONCOMITANT MEDICATIONS

For prior or concomitant medications, incomplete (i.e., partially missing) start date and/or stop date will be imputed. When the start date and the stop date are both incomplete for a patient, impute the start date first.

7.6.1 Incomplete Start Date

The following rules will be applied to impute the missing numerical fields. If the stop date is complete and the imputed start date is after the stop date, then the start date will be imputed using the stop date.

Missing day and month

- If the year of the incomplete start date is the same as the year of the date of the first dose of double-blind study drug, then the day and month of the date of the first dose will be assigned to the missing fields.
- If the year of the incomplete start date is prior to the year of the date of the first dose of double-blind study drug, then December 31 will be assigned to the missing fields.

Missing month only

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

Missing day only

- If the month and year of the incomplete start date are the same as the month and year of the date of the first dose of double-blind study drug, then the day of the date of the first dose will be assigned to the missing day.
- If either the year is before the year of the date of the first dose of double-blind study drug or if both years are the same but the month is before the month of the date of the first dose of double-blind study drug, then the last day of the month will be assigned to the missing day.

- If either the year is after the year of the date of the first dose of double-blind study drug or if both years are the same but the month is after the month of the date of the first dose of double-blind study drug, then the first day of the month will be assigned to the missing day.

7.6.2 Incomplete Stop Date

The following rules will be applied to impute the missing numerical fields. If the date of the last dose of double-blind study drug is missing, replace it with the last diary entry date. If the imputed stop date is before the start date (imputed or non-imputed start date), then the imputed stop date will be equal to the start date.

Missing day and month

- If the year of the incomplete stop date is prior to the year of the date of the last dose of double-blind study drug, then December 31 will be assigned to the missing fields.

Missing month only

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

Missing day only

- If either the year is before the year of the date of the last dose of double-blind study drug or if both years are the same but the month is before the month of the date of the last dose of double-blind study drug, then the last day of the month will be assigned to the missing day.

7.7 CHARACTER VALUES OF CLINICAL LABORATORY PARAMETERS

If the reported value of a clinical laboratory parameter cannot be used in a statistical summary table due to, for example, it being a character string rather than a numerical type, a coded value needs to be appropriately determined and used in the statistical analyses. However, the actual values as reported in the database will be presented in data listings.

Table 7.7-1 shows examples of how some possible laboratory results should be coded for the analysis.

Table 7.7-1. Examples of Coding Special Character Values for Clinical Laboratory Parameters

Laboratory Test, SI Unit	*Possible Laboratory Results*	*Coded Value for Analysis*
CHEMISTRY		
ALT, U/L	< 5	0
AST, U/L	< 5	0
Bilirubin, total, µmol/L	< 2	0

ALT = alanine aminotransferase; AST = aspartate aminotransferase; SI = *Le Système International d'Unités* (International System of Units).

7.8 HEALTH OUTCOMES PARAMETERS

7.8.1 Short Form 12 Health Survey Version 2

The SF-12 is a multipurpose, short-form health survey consisting of 12 questions designed for use in clinical practice and research, health policy evaluations, and general population surveys.

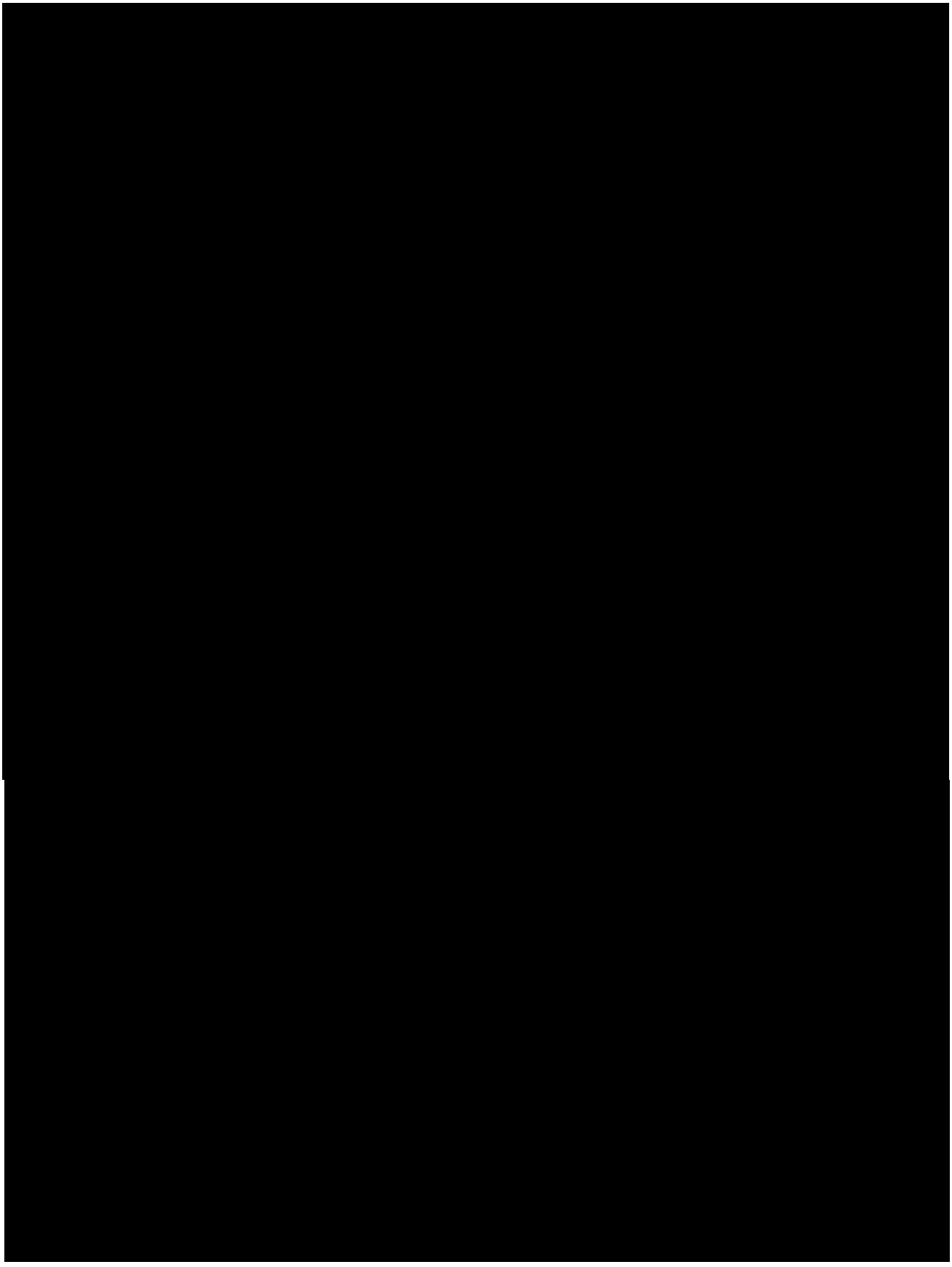
[redacted]

[redacted]

[redacted]

[redacted]

[redacted]



7.8.2 EuroQol-5 dimension (EQ-5D-3L)

The 5-item responses of the EQ-5D-3L define a health state that can be converted into a single summary utility index that assigns weights (ie, value) to each level. Different weights are available.

[redacted]

[redacted]

[redacted]

[redacted]

8. REFERENCES

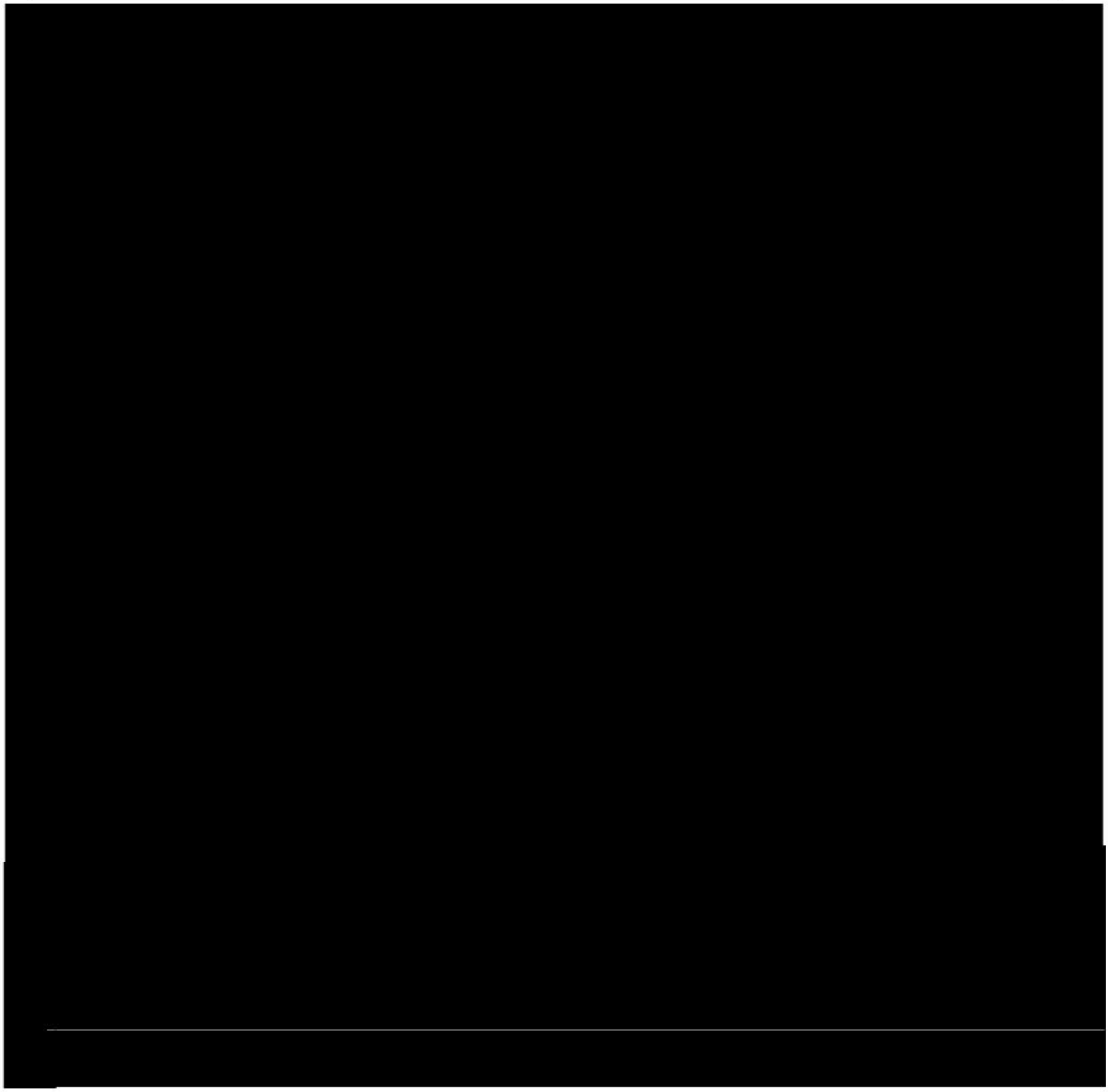
1. Dimenäs E, Glise H, Hallerbäck B, Hernqvist H, Svedlund J, Wiklund I. Quality of life in patients with upper gastrointestinal symptoms. An improved evaluation of treatment regimens? *Scand J Gastroenterol*.1993;28(8):681–687.

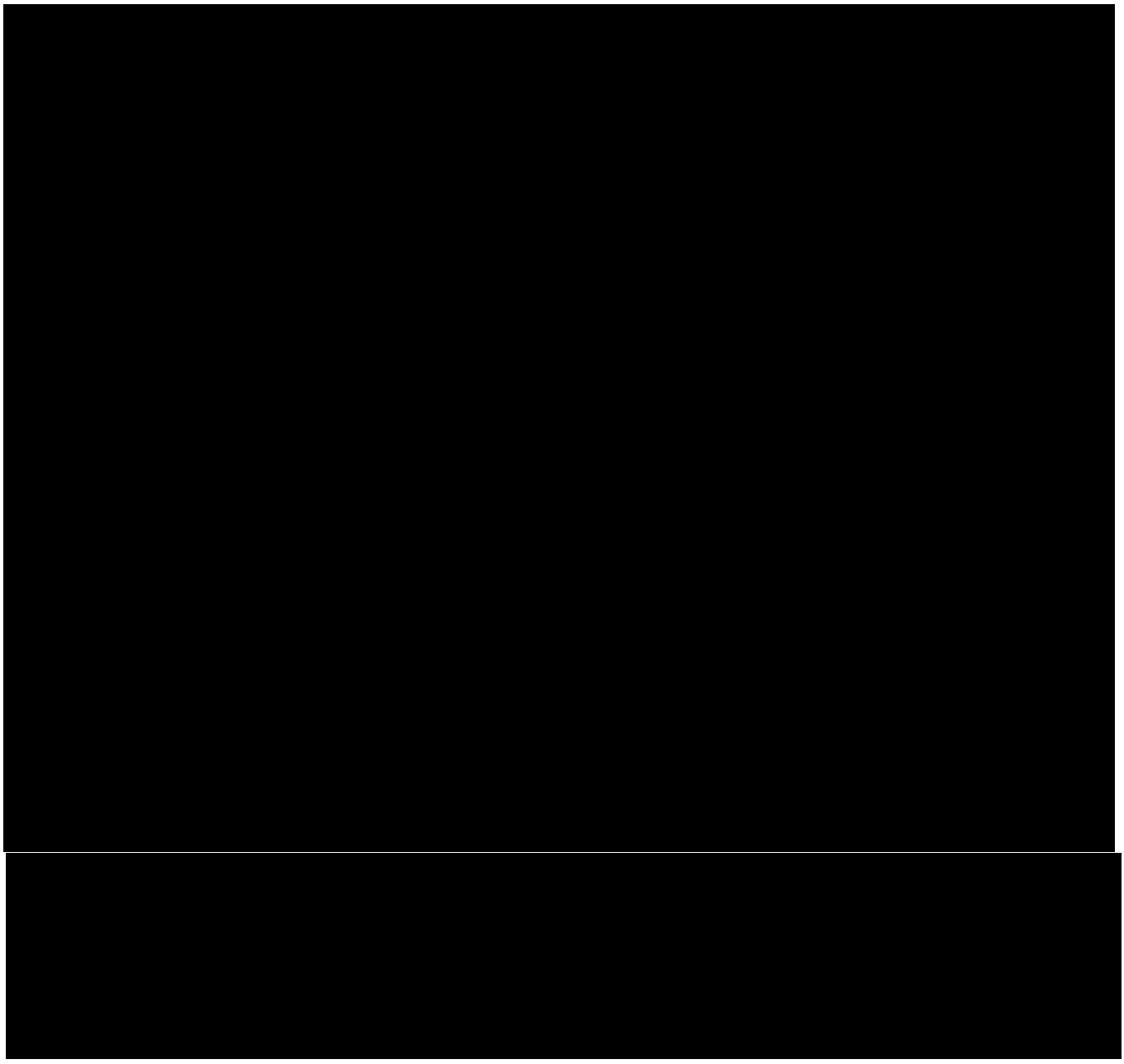
2. Ware John E.Jr., Kosinski M. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. Med Care 1996;34:220-33.

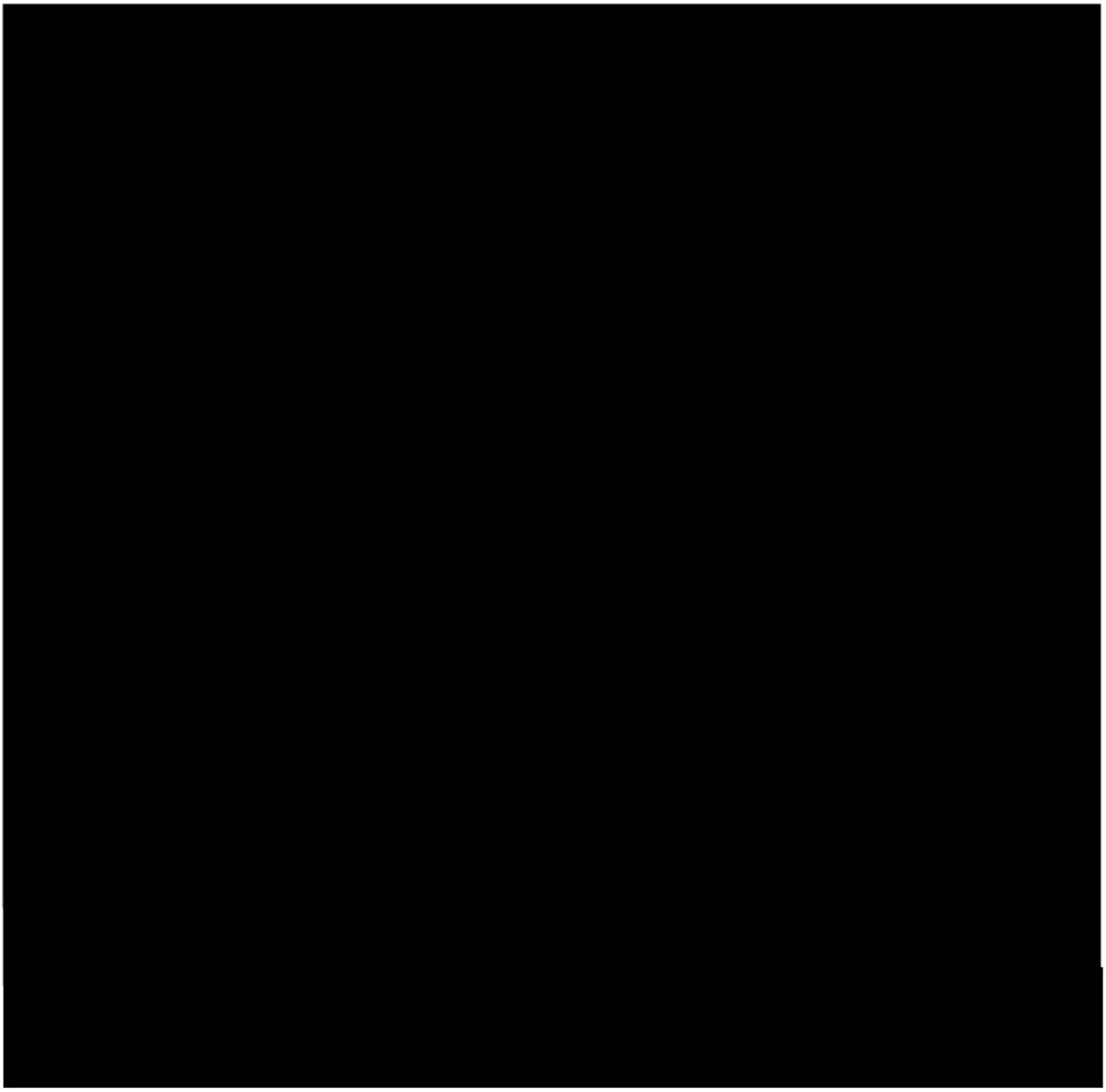
3. Shaw JW, Johnson JA, Coons SJ. U.S. valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care 2005;43:203-20.

9. APPENDICES

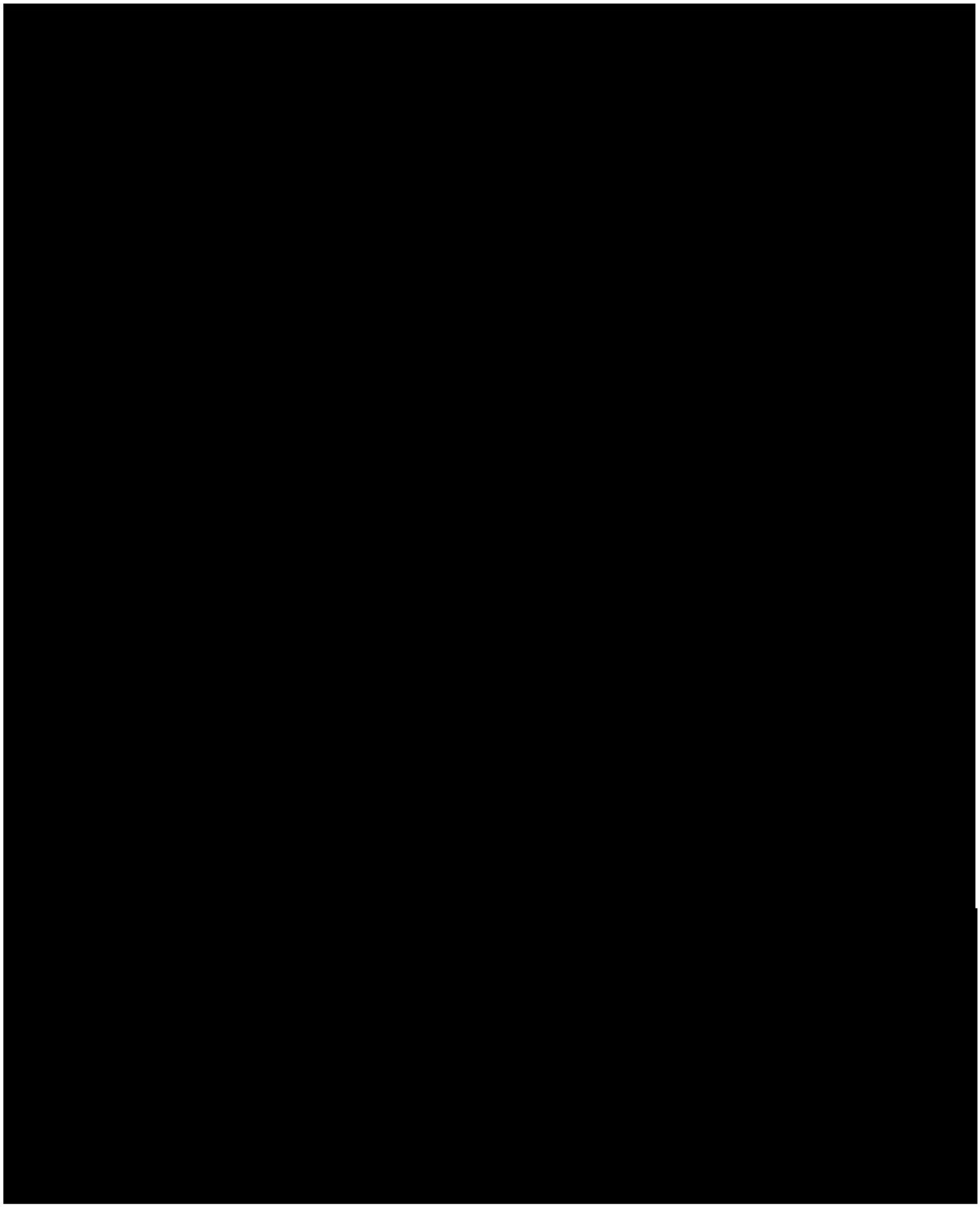
APPENDIX I GSRS-SELF

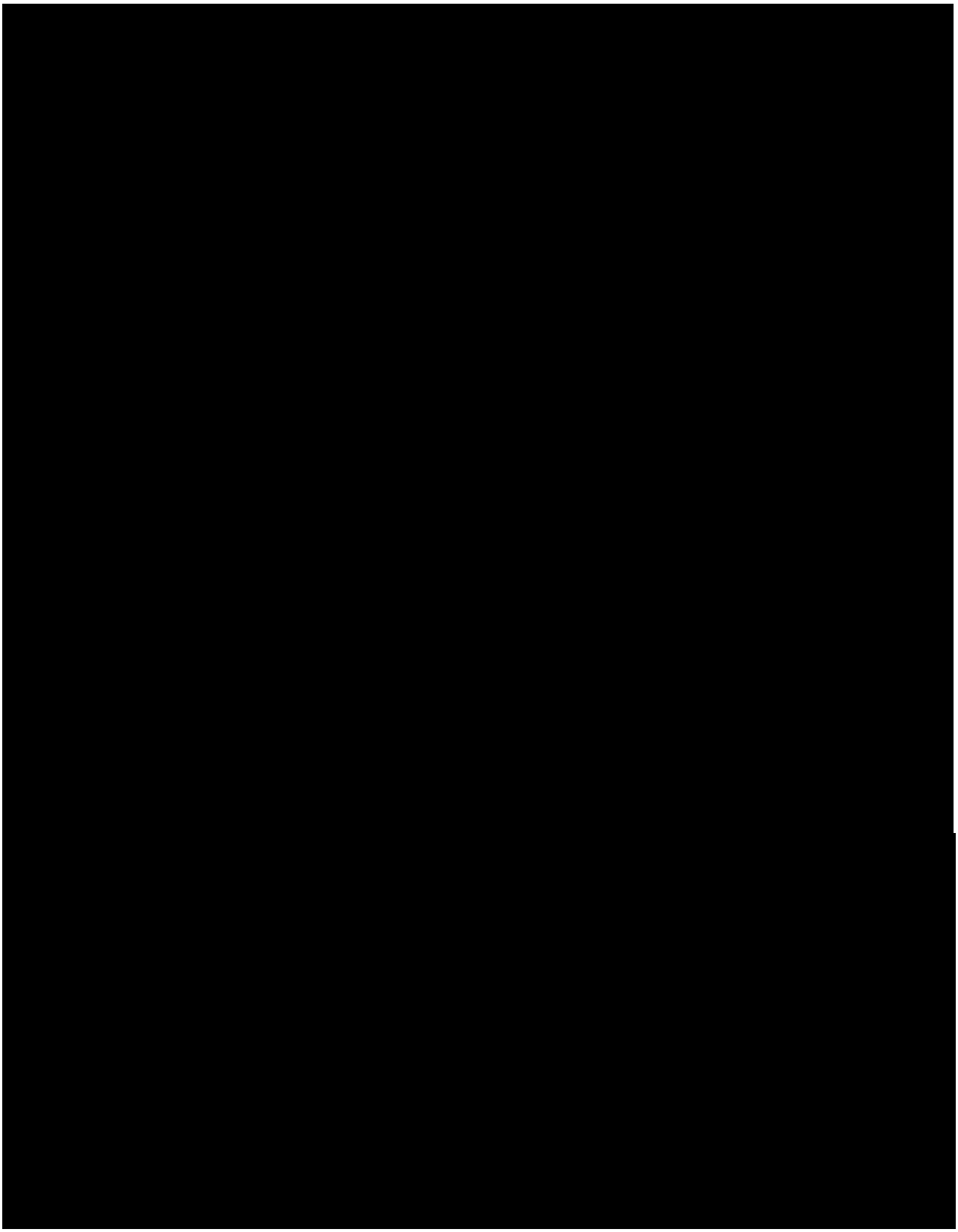


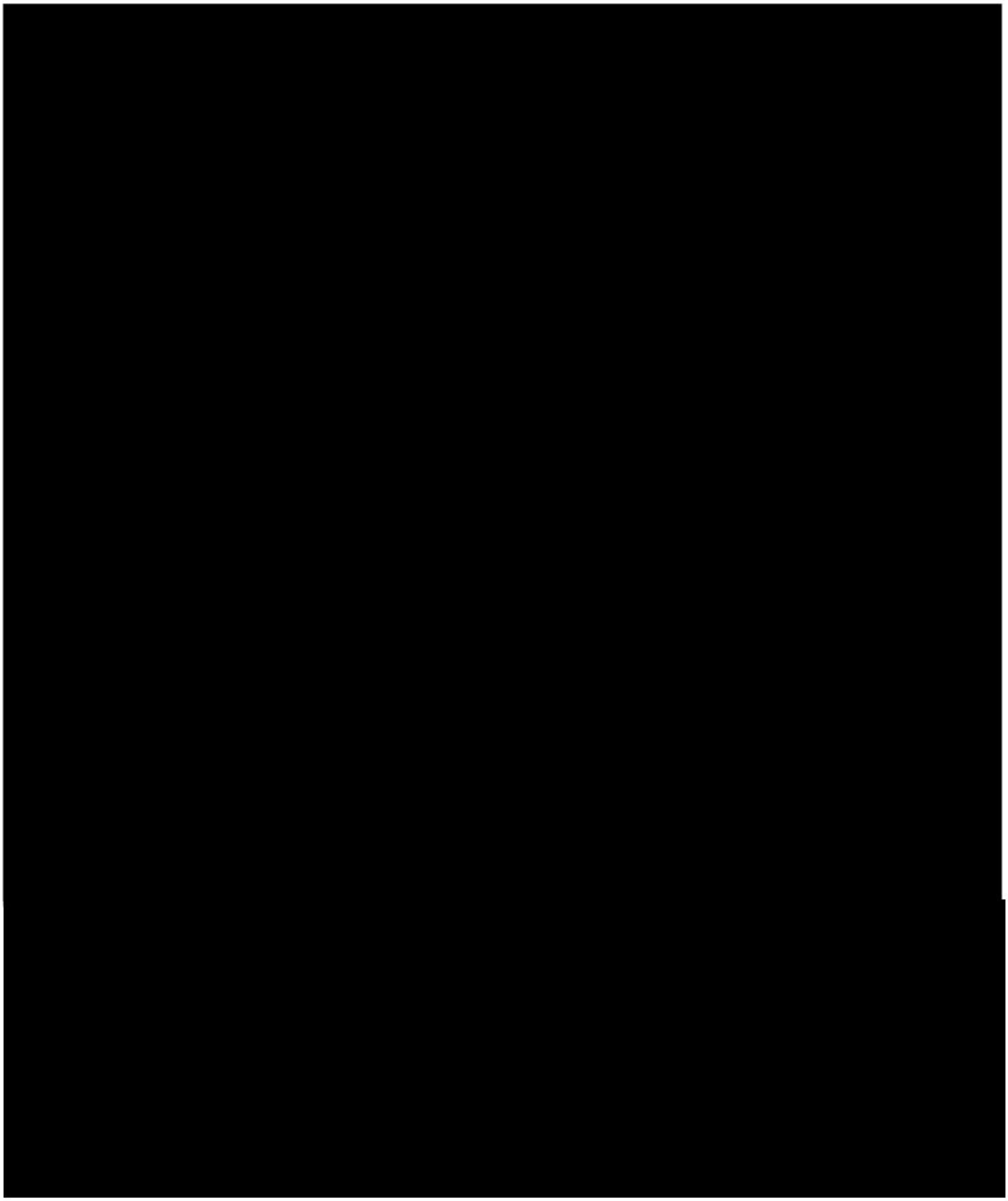




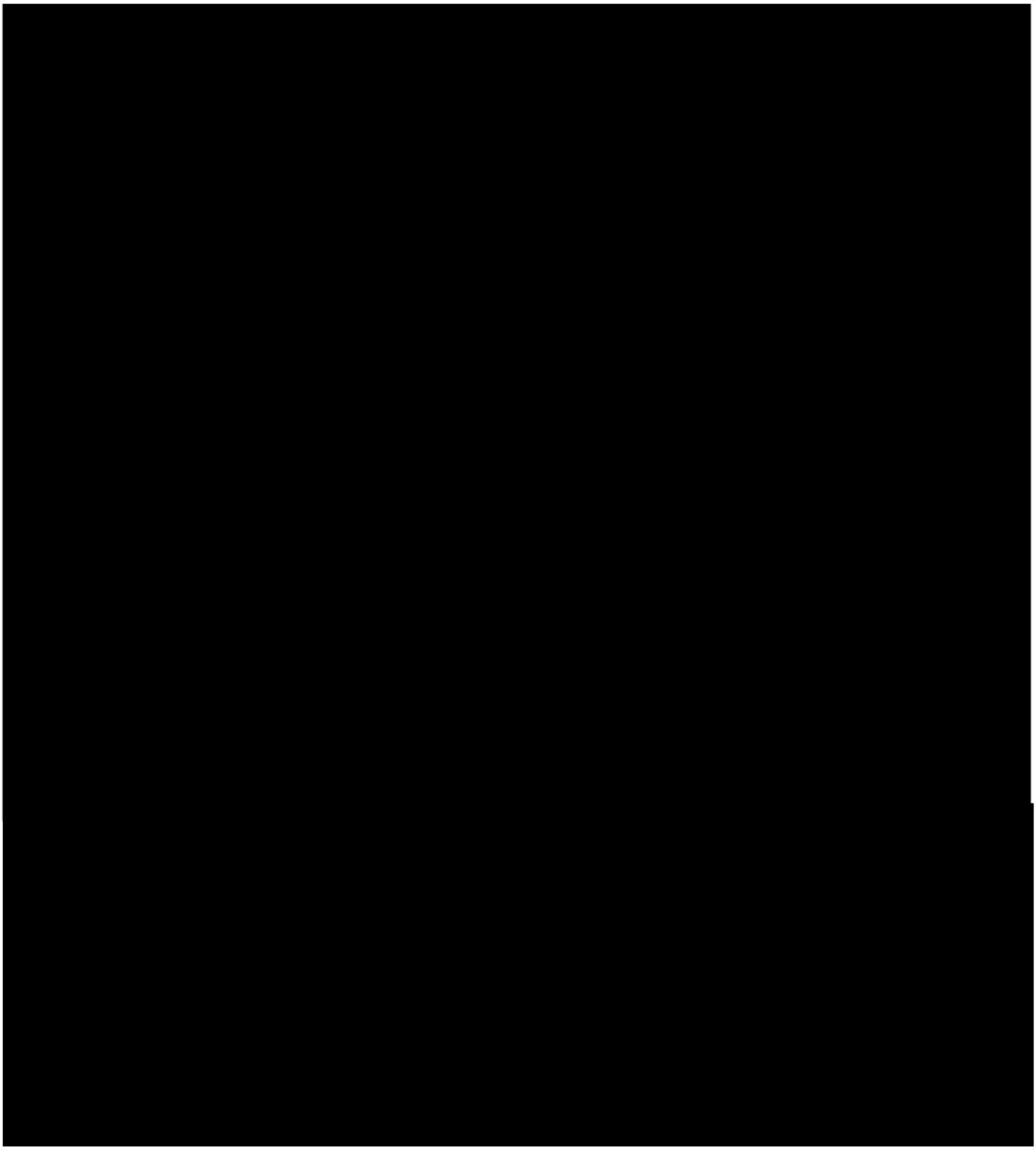
APPENDIX II SF-12V2







APPENDIX III EQ-5D-3L QUESTIONNAIRE



[redacted]