Statistical Analysis Plan (SAP)

Protocol Title:	A Randomized, Multicenter, Open-Label, Parallel-Group Clinical Study Comparing the Safety and Efficacy of MYL-1601D with NovoLog® in Type 1 Diabetes Mellitus Patients
Protocol Number:	MYL-1601D-3001
Protocol Version, Date	Version 5.0, 14 October 2019
ICON ID:	Mylan 0102/0045
Document Version, Date:	3.0, 02 March 2020

Prepared by:

ICON Clinical Research Services

On behalf of:

Mylan GmbH Pharmaceuticals Inc.

SIGNATURE PAGE

Prepared at ICON Clinical Research by:

Xia Deng
Xia Deng, Senior Biostatistician I

02 MAR 2020
Date
(DD MMM YYYY)

Reviewed at ICON Clinical Research by:

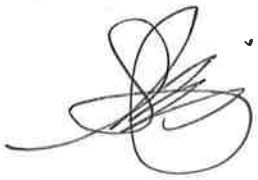
Wenjun Xin
Wenjun Xin, Senior Biostatistician I

02MAR2020
Date
(DD MMM YYYY)

Charissa Hogeboom, Senior Program Manager

02 Mar 2020
Date
(DD MMM YYYY)

Approved at Mylan Pharmaceutical Inc by:



Bin Sun, Principal Biostatistician

02 Mar 2020

Date
(DD MMM YYYY)

REVISION HISTORY

Version/Date	Version name	Section	Changes implemented
FINAL 1.0/11JAN2019	Final Version 1.0	Not Applicable	Not Applicable
FINAL 2.0/18OCT2019	Final Version 2.0	7.2.1 8.1 8.7.1 8.7.2 8.7.3.1 8.7.5 8.8.4 8.8.5 8.9 10.1 10.2 10.3	Updated missing data analysis method for primary analysis due to immunogenicity assay method changes Updated imputation for retests, unscheduled visits and visit end of treatment; updated visit window Updated analysis methods for primary endpoint Updated primary efficacy analysis method and tipping point analysis due to immunogenicity assay method changes Added sensitivity analysis for HbA1c Updated subgroup analysis method Updated the definition for relative hypoglycemia Updated Immunogenicity assessments Updated the incidence analyses for lab data, vital signs, physical examination and ECG Updated multiple imputation method SAS code, added multiple imputation for HbA1c Updated example SAS code for tipping point analysis method Updated SI units for laboratory category table
FINAL V3.0/ 02MAR2020		6.1 6.2 6.3	Add randomized analysis set definition. Modified ITT analysis set to exclude subjects who are in the site with GCP violation Modified PP analysis set to exclude subjects who are in the site with GCP violation and also add more clarification for PP analysis set.

		6.4	Modified Safety analysis set to exclude subjects who are in the site with GCP violation
		7.2	Change single imputation into multiple imputations for missing baseline value.
		7.3	Add a section address GCP issue at one investigator site (Site ID: 6103) and add sensitivity analysis without inclusion of data from this site
		10.0	Add SAS programming flow chart per regulatory request

TABLE OF CONTENTS

SIGNATURE PAGE 2
REVISION HISTORY 4
TABLE OF CONTENTS 6
LIST OF ABBREVIATIONS 8
1 INTRODUCTION 10
2 STUDY OBJECTIVES 11
2.1 Primary objective(s) 11
2.2 Other objective(s) 11
3 STUDY DESIGN 12
3.1 General study design 12
3.2 Randomization and blinding 12
3.2.1 Randomization 12
3.2.2 Breaking the blinding and study team blindness 13
3.3 Study treatments and assessments 13
4 STUDY ENDPOINTS 20
4.1 Primary efficacy endpoint(s) 20
4.2 Secondary efficacy endpoint(s) 20
4.3 Safety endpoint(s) 20
4.4 Exploratory endpoint(s) 20
5 SAMPLE SIZE AND POWER 21
6 ANALYSIS POPULATIONS 22
6.1 Randomized analysis set 22
6.2 ITT analysis set 22
6.3 PP analysis set 22
6.4 Safety analysis set 22
6.5 Run-in analysis set 22
6.6 Protocol deviations/violations and exclusions from analysis sets 22
7 Statistical Considerations and analysis 23
7.1 Derived Variables 23
7.2 Handling of missing data and outliers 23
7.2.1 Missing data analysis methods 23
7.2.2 Handling of missing weight 25
7.2.3 Handling of missing or incomplete dates 25
7.3 Handling of site with GCP violation 26
8 STATISTICAL METHODS 27
8.1 General statistical conventions 27
8.2 Investigator pooling for secondary analyses 28
8.3 Subject disposition 28
8.4 Protocol deviations 28
8.5 Demographics and baseline characteristics 29
8.5.1 Demographics 29
8.5.2 Baseline and disease characteristics 29

8.5.3 Medical history 29
8.5.4 Concomitant medications 29
8.6 Extent of exposure 30
8.6.1 Treatment duration 30
8.6.2 Treatment compliance 30
8.7 Efficacy analyses 30
8.7.1 Analysis methods 30
8.7.1.1 Multiplicity 31
8.7.1.2 Treatment by center interaction analysis (multi-center study) 31
8.7.2 Analysis of primary efficacy endpoint(s) 31
8.7.2.1 Primary efficacy analysis 31
8.7.2.2 Secondary/Sensitivity analyses of the primary efficacy endpoint 34
8.7.2.3 Tipping point analysis 35
8.7.3 Secondary endpoints 35
8.7.3.1 HbA1c 36
8.7.3.2 Fasting plasma glucose 36
8.7.3.3 Daily insulin dose per unit body weight (mealtime, basal insulin and total) for days of 7-Point profiles 37
8.7.3.4 7-Point Self-Monitored Blood Glucose (SMBG) profile 38
8.7.4 Analysis of exploratory endpoint(s) 39
8.7.5 Subgroup analyses 39
8.8 Safety analyses 40
8.8.1 Adverse events 41
8.8.2 Device-related Assessment include both TEAE and device complaints 42
8.8.3 Hypoglycemia 43
8.8.4 Hypersensitivity and Immune mediated adverse events 44
8.8.5 Immunogenicity assessments 44
8.9 Other Safety Analyses 45
8.9.1 Laboratory Data 45
8.9.2 Vital signs 46
8.9.3 Physical examinations 46
8.9.4 Electrocardiograms (ECG) 47
8.10 Interim analysis 47
9 REFERENCES 48
10 APPENDICES – SAS example codes 49
10.1 Multiple imputation 49
10.1.1 Multiple imputation for binary ADA response if baseline ADA is negative. 51
10.1.2 Multiple imputation for continuous ADA titer if baseline ADA positive. 56
10.1.3 Combining the multiple imputation results to derive TEAR status 58
10.1.4 Multiple imputation for HbA1c. 60
10.2 Tipping Point Analysis 63
10.3 Laboratory Category 68

LIST OF ABBREVIATIONS

The following abbreviations will be used within this SAP.

ADA	Anti-Drug Antibody
AE	Adverse Event
ALT	Alanine Transaminase
AST	Aspartate Transaminase
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
CI	Confidence Interval
CM	Concomitant Medication
CRF	Case Report Form
CRO	Contract Research Organization
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria Adverse Events
CV	Coefficient of Variation
DM	Diabetes Mellitus
ECG	Electrocardiogram
FDA	Food and Drug Administration
FPG	Fasting Plasma Glucose
GCP	Good Clinical Practice
HbA1c	Glycosylated hemoglobin
HBsAg	Hepatitis B Surface Antigen
HCVAb	Hepatitis C antibodies
HIV	Human Immunodeficiency Virus
HR	Heart Rate
ICH	International Council for Harmonisation
IXRS	Interactive Web/Voice Response System
ITT	Intent to treat
IU	International Unit
Kg	Kilogram
MCMC	Markov Chain Monte Carlo
MedDRA	Medical dictionary for regulatory activities
mL	Milliliter
MMRM	Mixed Model Repeated Measures

NAb	Neutralizing Antibody
NCI	National Cancer Institute
PD	Pharmacodynamic
PP	Per Protocol
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SID	Subject Identification
SMBG	Self-monitored blood glucose
SOC	System Organ Class
SOP	Standard Operating Procedure
T1DM	Type 1 diabetes mellitus
TEAR	Treatment Emergent Antibody Response
TEAE	Treatment Emergent Adverse Event
WHODDE	World Health Organization Drug Dictionary Enhanced

1 INTRODUCTION

The purpose of this Statistical Analysis Plan (SAP) is to provide detailed descriptions of the statistical methods, data derivations and data displays for study protocol MYL-1601D-3001 Version 5.0 "A Randomized, Multicenter, Open-Label, Parallel-Group Clinical Study Comparing the Safety and Efficacy of MYL-1601D with NovoLog® in Type 1 Diabetes Mellitus Patients" dated 14 October 2019 for Clinical Study Report (CSR) analysis and interim analyzes, if any. The table of contents and templates for the tables, figures and listings (TFLs) will be produced in a separate document.

Any deviations from this SAP will be described and justified in the CSR.

The preparation of this SAP has been based on International Conference on Harmonisation (ICH) E9 guidelines.

All data analyzes and generation of TFLs will be performed using SAS 9.3® or higher.

2 STUDY OBJECTIVES

2.1 Primary objective(s)

The primary objectives of the study are as follows:

To demonstrate that immunogenicity as assessed by treatment emergent antibody response (TEAR) rate with MYL-1601D is equivalent to that of NovoLog® during 24-week treatment. The TEAR rate will not be assessed in isolation but will be part of totality of evidence including changes in HbA1c, fasting plasma glucose (FPG), insulin dose, neutralizing antibodies and injection site reactions to ensure that changes in TEAR rate, if any, are clinically meaningful.

2.2 Other objective(s)

The other objectives of the study is to compare MYL-1601D to NovoLog® administered in combination with Lantus®, with respect to:

- Immunogenicity assessments: visits TEAR assessments, incidences of Anti-Drug Antibody (ADA) and positive Cross-reactive ADA, and positive Neutralizing Antibody (NAb)
- Analysis of impact of ADA on pharmacodynamic (PD) parameters such as FPG, HbA1c, and insulin dose.
- Incidence and rate of hypoglycemic events
- Occurrence of local reactions, systemic reactions and other adverse events
- Safety related to hypersensitivity and immune mediated adverse events (subgroup analysis)
- Change in HbA1c from baseline
- Change in fasting plasma glucose (FPG) from baseline
- Change in prandial, basal, and total daily insulin dose per unit body weight (U/kg) from baseline
- Change in 7-point SMBG profile from baseline
- Device-related safety assessment

3 STUDY DESIGN

3.1 General study design

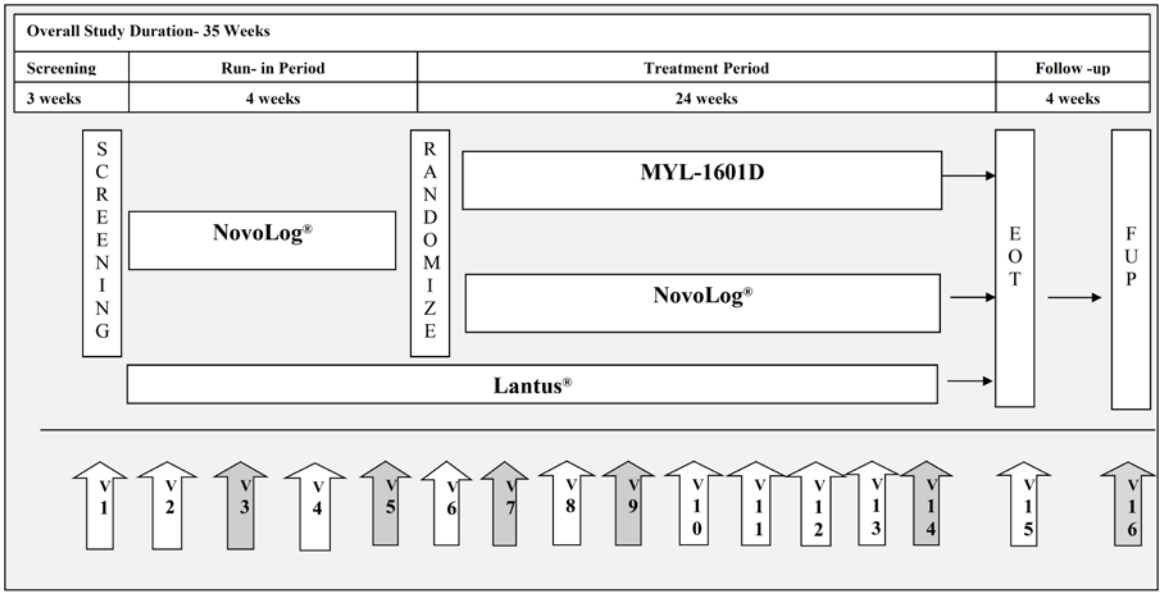
This is a multicenter, open-label, randomized, parallel-group phase 3 study in subjects with Type 1 Diabetes Mellitus (T1DM) comparing the safety and efficacy of MYL-1601D with NovoLog®.

After up to 3-week screening period, all subjects will be titrated on NovoLog® during a 4-week run-in period, and will be shifted from their current basal insulin to study insulin Lantus®. After run-in period, subjects will be randomized; one group will receive MYL-1601D, while the other group will receive NovoLog® for 24 weeks. A follow-up visit, via telephone call, will be scheduled 4 weeks after last dose of MYL-1601D.

The study will be conducted at approximately 200 sites in the United States (US).

The Study Flow Chart is presented in Figure 1.

Figure 1: Study Flow Chart



Abbreviations: V=Visit; EOT=End of Treatment; FUP=Follow-up;

3.2 Randomization and blinding

3.2.1 Randomization

Assignment of Subject Identification number (SID), randomization number and study medication, as well as site drug inventory control will be managed by an automated IWRS. A manual containing complete instructions for Web or telephone access and use will be provided to each site prior to study start. The IWRS will assign a SID for each subject's first clinic visit. Each SID will be unique and serve as the primary subject identifier

throughout all phases of the study. The SID must appear on all case report form (CRF) pages, source documents, laboratory data, ECG and diary data. Subjects qualifying to enter the study drug treatment phase, will be assigned an additional "randomization number" by the IWRS at randomization. Dynamic allocation with minimization algorithm will be used for treatment randomization. Randomization will be stratified by investigator and basal insulin (Glargine) dose time (morning or evening).

3.2.2 Breaking the blinding and study team blindness

The trial design is open-label since the two products, MYL-1601D and NovoLog®, have distinct packaging. To avoid any bias in the evaluation of the critical endpoints, a blinded analysis of immunogenicity and other parameters such as HbA1c and FPG is planned. In addition, most of the study personnel (Sponsor and CRO teams) will be blinded to the randomized assigned treatment arms. This will be detailed in separate document, listing the Sponsor and CRO personnel who are unblinded to ensure transparency and proper study conduct. As all investigators and sites staff are not blinded to treatment, no process is required for breaking the blind.

3.3 Study treatments and assessments

During the Run-in period, all subjects will receive FlexPen NovoLog® from Novo-Nordisk (US listed drug) 100 U/mL until randomization. In addition, all patients will be shifted from their current basal insulin to Lantus® SoloSTAR® at the start of the run-in period, and will continue this for the complete study duration. The doses of NovoLog® and Lantus® will be titrated (if required) during the run-in period to ensure diabetes control.

During the Treatment period, all patients will receive one of the following treatments:

- MYL-1601D or FlexPen NovoLog® to be taken at meal time. Both investigational products will be provided in a pre-filled disposable pen with a 3-mL cartridge. During the treatment period, dose titration will be kept to a minimum.
- Once daily Lantus® SoloSTAR® (insulin glargine injection, 100 U/mL), manufactured by Sanofi-Aventis.

A detailed description of procedures and assessments to be conducted during this study is summarized in the Scheduled of Study Assessments in Table 1 below.

Table 1: Schedule of Study Assessments

Study Periods	Screening	Run-in Period				Randomized Comparative Treatment Period										Follow-up
Study Visits [1]	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15 (EOT) [10]	V16 (FU)
Study Week	-7 to -4	-4	-3	-2	-1	0	1	2	4	8	12	16	20	22	24	28
Study Days	-49 to -28	-28±3	-21±3	-14±3	-7±3	0±3	7±3	14±3	28±3	56±7	84±7	112±7	140±7	154±7	168±7	196±7
Informed Consent	x															
Inclusion/Exclusion Criteria Review	x					x										
History of previous insulin usage	x															
Dilated Ophthalmoscopy / retinal photography (if not done the last 6 months)	x															
Standard-of-care specifics [2]	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
Age, Gender, Height, Race	x															

Study Periods	Screening	Run-in Period				Randomized Comparative Treatment Period										Follow-up
Study Visits [1]	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15 (EOT) [10]	V16 (FU)
Study Week	-7 to -4	-4	-3	-2	-1	0	1	2	4	8	12	16	20	22	24	28
Study Days	-49 to -28	-28±3	-21±3	-14±3	-7±3	0±3	7±3	14±3	28±3	56±7	84±7	112±7	140±7	154±7	168±7	196±7
Body Weight and BMI [11]	x	x		x		x		x	x	x	x	x	x		x	
Pregnancy Test [3]	x	x				x			x	x	x	x	x		x	
Medical History and concomitant illness	x															
Concomitant Medications	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Vitals signs measurement (sitting)	x	x		x		x		x	x	x	x	x	x		x	
Physical examination	x					x									x	
12-lead ECG (supine)	x														x	

Study Periods	Screening	Run-in Period				Randomized Comparative Treatment Period										Follow-up
Study Visits [1]	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15 (EOT) [10]	V16 (FU)
Study Week	-7 to -4	-4	-3	-2	-1	0	1	2	4	8	12	16	20	22	24	28
Study Days	-49 to -28	-28±3	-21±3	-14±3	-7±3	0±3	7±3	14±3	28±3	56±7	84±7	112±7	140±7	154±7	168±7	196±7
Randomization						x [8]										
Record AEs and SAEs (including local and systemic allergic reactions) and hypoglycemic events [4] due to medication, disposable pen or needle	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Record device complaint			x	x	x	x	x	x	x	x	x	x	x	x	x	
Fasting plasma glucose	x					x			x	x	x	x	x		x	
HbA1c Assay	x					x					x				x	
Rescue criterion evaluation [6]												x				

Study Periods	Screening	Run-in Period				Randomized Comparative Treatment Period										Follow-up
Study Visits [1]	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15 (EOT) [10]	V16 (FU)
Study Week	-7 to -4	-4	-3	-2	-1	0	1	2	4	8	12	16	20	22	24	28
Study Days	-49 to -28	-28±3	-21±3	-14±3	-7±3	0±3	7±3	14±3	28±3	56±7	84±7	112±7	140±7	154±7	168±7	196±7
Fasting C-peptide, HIV, HBsAg, HCVAb	x															
Sampling for hematology, blood chemistry and urinalysis [5]	x					x									x	
Fasting lipid profile						x									x	
Sampling for immunogenicity [9]	x					x		x	x	x	x	x	x		x	
Review diary, 7-point SMBG Profile and dose collection performed in the week before the visit [7]				x		x		x	x	x	x	x	x		x	
Dose review of MYL-1601D, NovoLog® and Lantus® and dose		x	x	x	x	x	x	x	x	x	x	x	x	x	x	

Study Periods	Screening	Run-in Period				Randomized Comparative Treatment Period										Follow-up
Study Visits [1]	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15 (EOT) [10]	V16 (FU)
Study Week	-7 to -4	-4	-3	-2	-1	0	1	2	4	8	12	16	20	22	24	28
Study Days	-49 to -28	-28±3	-21±3	-14±3	-7±3	0±3	7±3	14±3	28±3	56±7	84±7	112±7	140±7	154±7	168±7	196±7
adjust/instruction																
Dispense study medication and ancillary supplies		x				x			x	x	x	x	x			
Drug Accountability and Compliance				x		x		x	x	x	x	x	x		x	
Dispense subject diary		x		x		x		x	x	x	x	x	x			

1. Visits 3, 5, 7 14 and 16 may be telephone contacts/visits, in case required an actual site visit is possible (grey columns represent telephone contacts).
2. Standard–of-care specifics includes assessment and documentation of the following - Training on self-management of diabetes, lifestyle modification measures (includes maintenance of appropriate body weight, following recommended physical activity, avoidance of smoking and following the recommended diet); and monitoring to prevent complications.
3. Serum pregnancy test for women of child bearing potential will be done during screening and randomization visits (V1 and V6). During subsequent visits urine pregnancy test will be done, any positive urine test needs to be confirmed with serum test. At the randomization visit, both urine and serum pregnancy tests will be done, subject can be enrolled only if the urine pregnancy test is negative until serum result is provided.
4. Non-severe hypoglycemic events (which are not consider as SAE) occurring after the EOT visit will not be recorded at the follow-up visit.
5. A routine urine dipstick test will be performed by the site. A microscopic urinalysis will be performed by the central lab if the dipstick test result is abnormal and the Investigator deems it clinically significant and requests further evaluation.
6. Rescue criterion is evaluated on V12/Week 16 based on HbA1c measurement at V11/Week 12. Sites are required to receive the V11/Week 12 results prior to V12/Week 16 subject visit to enable Investigator to take decision on future steps.
7. The 7-point SMBG profile recorded via glucometer and recorded by the subject in the diary, measurement will be performed by the subject at home on 3 days (of which 2 days are consecutive) in the week before the visit.
8. Prior to randomization, Investigator is required to re-confirm subject eligibility to the study based on the data collected during the screening period, including the labs values recorded during screening.
9. Immunogenicity samples: At Visits 1, 6, 8, 9, 10, 11, 12, 13 and 15. At each visit, five blood samples of 5 mL each (25 ml per visit) will be drawn for ADA and NAb.
10. In case the subject discontinues treatment early, the site should explain the importance of data collection and make every effort to retain the subject, perform the remaining visits, off study treatment, per the Schedule of Study Assessments, up to the completion of Week 28/Follow-up call. BMI is calculated at screening visit, randomization visit, week 12 and week 24.

4 STUDY ENDPOINTS

4.1 Primary efficacy endpoint(s)

The TEAR rate during 24-week treatment period is the primary endpoint for this study.

TEAR is defined as either one of the following:

1) Subjects who are ADA negative at baseline and become positive at any timepoint post baseline

2) Subjects who are ADA positive at baseline and demonstrate 4-fold increase in titer values at any timepoint post baseline visit.

4.2 Secondary efficacy endpoint(s)

The secondary endpoint points are change from baseline to week 24 for the ITT population.

- Change in HbA1c from baseline
- Change in fasting plasma glucose from baseline
- Change in prandial, basal and total daily insulin dose per unit body weight (U/kg) from baseline
- Change in 7-point SMBG profile from baseline

4.3 Safety endpoint(s)

- Incidence of positive antibody response and NAb and change in antibody percentage binding from baseline.
- Impact of ADA on PD parameters, such as FPG, HbA1c, and insulin dose.
- Change in hypoglycemia rate (30 day adjusted) from baseline and incidence of hypoglycemic events.
- Incidence of treatment-emergent adverse events (TEAEs) and serious adverse events (SAEs).
- Incidence of local reactions (includes injection site reaction), systemic reactions.
- Incidence of hypersensitivity and immune mediated adverse events.
- Incidence of device-related safety assessment.

4.4 Exploratory endpoint(s)

Not Applicable.

5 SAMPLE SIZE AND POWER

Approximately 500 subjects with type 1 diabetes will participate in this trial. The sample size estimation is based on primary objective - to demonstrate that immunogenicity as assessed by TEAR rate with MYL-1601D is equivalent to that of NovoLog® during 24-week treatment. The 80% power will be achieved with 250 subjects per treatment arm to demonstrate that 90% confidence interval of treatment difference (MYL-1601D minus NovoLog®) is within prespecified ±margin (Table 2 of the SAP) (Chow SC et al, 2002; Chow S-C et al 2003). The pre-specified equivalent margin is dependent on the final rate of NovoLog® TEAR rate. No replacement of subject will be performed if subject discontinued prematurely from the study.

Table 2: Margins and 95% CIs with different reference TEAR event rates in 500 subjects

Event Rate for Reference Product	Estimated 95% CI for the Event Rate (n=250)	Margin	Type I error 2 one-sided alpha	Power	Total N
5%	(2%,8%)	5.7%	0.05	80%	500
10%	(6%,14%)	7.9%	0.05	80%	500
15%	(11%,19%)	9.3%	0.05	80%	500
20%	(15%,25%)	10.5%	0.05	80%	500
25%	(20%,30%)	11.3%	0.05	80%	500
30%	(24%,36%)	12.0%	0.05	80%	500
35%	(29%,41%)	12.5%	0.05	80%	500
40%	(34%,46%)	12.8%	0.05	80%	500
45%	(39%,51%)	13.0%	0.05	80%	500
50%	(44%,56%)	13.1%	0.05	80%	500
55%	(49%,61%)	13.0%	0.05	80%	500
60%	(54%,66%)	12.8%	0.05	80%	500

6 ANALYSIS POPULATIONS

6.1 Randomized analysis set

Randomized analysis set include all subjects who are randomized into the study.

6.2 ITT analysis set

The ITT analysis set includes all randomized subjects without being in the site with GCP violation (including subjects who receive incorrect treatment, do not complete the study or do not comply with the protocol or used prohibited medication, randomized but do not take any study drug). The subjects who are in the site with GCP violation will be excluded from ITT analysis set.

6.3 PP analysis set

The PP analysis set includes subjects who complete Week 24, and do not have protocol violations that impact the primary outcome (as detailed in the blind data review (BDR) plan). Subjects who take other fast acting insulin other than assigned study medication such as rescue medication or subject's own insulin will be excluded from PP analysis set. Subjects who are in the site with GCP violation will also be excluded from PP analysis set. The subjects excluded from the PP population will be determined during the blind data review meeting and summarized in the blind data review report.

6.4 Safety analysis set

The safety analysis set includes subjects who take at least one dose of the study medication after randomization and who aren't in the site with GCP violation. For safety analyses, subjects will be categorized according to the treatment that they actually received.

6.5 Run-in analysis set

Run-in analysis set includes all subjects who enrolled into run-in period.

6.6 Protocol deviations/violations and exclusions from analysis sets

All violations and exclusions of subjects from analysis sets will be identified at the Classification Meeting just prior to study unblinding, through clinical review input provided by Sponsor, using the following sources of information:

- Supportive subject listings, provided by the ICON lead statistician ahead of the Classification Meeting, based on data recorded on the electronic case report form (eCRF).
- Protocol Deviation Logs, provided by ICON Medical.

Further, deviations from the protocol will be classified as major or minor.

7 STATISTICAL CONSIDERATIONS AND ANALYSIS

7.1 Derived Variables

The below table provides the list of derived variables for Demographic and baseline characteristics, various duration derivations, drug compliance, baseline derivations and other important derivations applicable for this study.

Table 3.

Variables	Formula
Demographic and Baseline characteristics	
Body mass index (BMI) (kg/m2)	weight (kg) / [height (m)]^2
Derivation of Duration	
Study day at any visit	Date of interest – date of first dose of study drug. One day is added if this difference is ≥ 0
Extent of Exposure (Days)	Date of last randomized study medication intake – Date of first randomized study medication intake + 1
Extent of Exposure (Weeks)	Extent of exposure (days)/7
Baseline Derivations	
Baseline	Last non-missing value collected prior to the first dose of study medication in treatment period.
Change from baseline	Post baseline value – Baseline
SMBG Derivations	
SMBG excursion	Post-meal blood glucose – Pre-meal blood glucose
Mean SMBG	All BG/7 and average over 3 days
Mean excursion	All excursion values average over 3 days

7.2 Handling of missing data and outliers

7.2.1 Missing data analysis methods

Missing data will only be imputed in the primary and sensitivity analyses for TEAR rate associated with ADA and other secondary variables that are mentioned in this document. Otherwise, missing data will not be imputed.

Subjects with missing data are either 1) subjects who completely miss post-baseline ADA data or 2) subjects who miss a portion of ADA data and for whom the TEAR positive or negative status cannot be determined. If a subject with partially missing ADA data already

meets the TEAR criteria, the subject is not considered to be a subject with missing data.

To minimize any bias due to missing data, multiple imputation will be used to fill in missing values according to either of two TEAR criteria (binary response and continuous titer values) prior to deriving the TEAR and subsequently estimating the treatment difference and 90% confidence interval. Two imputation processes will be used for each of two TEAR criteria among the missing data. For baseline ADA negative subjects, post baseline binary response (positive or negative) will be imputed using logistic regression multiple imputation assuming missing not at random. For baseline positive subjects, missing titer values (continuous values) will be imputed with same treatment group non-missing subjects using pattern mixture model with the complete-case missing values (CCMV) method (Little 1993). In the case where there are sufficient number of retrieved dropouts, that is, if there are 6 retrieved dropout subjects in each treatment group, the missing values in that treatment group will be imputed by modelling the retrieved subgroup data. Retrieved dropout subjects are the subjects who discontinued assigned study treatment but still remain on the study and had their efficacy and safety measurements captured at the planned visits. If less than 6 retrieved dropout subjects in each treatment group, all non-missing subjects within same treatment group will be used to impute missing data by multiple imputation.

If baseline value is missing, the missing baseline value will be imputed with multiple imputations (6 imputations) using no missing baseline values from same treatment group.

After imputation, the TEAR criteria will then be applied to each imputed dataset to obtain if a subject is TEAR positive or negative, and the treatment difference and 90% confidence interval will be estimated using Wald confidence limit method. The result will be combined over all imputations by SAS PROC MIANALYZE.

Tipping-point analysis will be performed via the method proposed by Liublinska and Rubin (2014). This will include a graphic display to indicate the tipping point analysis for the TEAR status binary outcome. A graphic display (heating map) is used to indicate the tipping point analysis results with missing values as TEAR positive. The treatment differences are presented as values in the heat map. The graph will be displayed to show the sensitivity ranging from missing favour testing treatment to missing favour reference treatment.

For tipping-point analysis, the number of subjects with missing data include 1) subjects who are missing all post-baseline ADA data, 2) subjects with partially missing ADA data whose TEAR status cannot be determined as positive or negative, and 3) subjects who are missing ADA data at baseline. If a subject with partially missing ADA data already meets the TEAR criteria, the subject is not considered as a subject with missing data.

7.2.2 Handling of missing weight

In the secondary efficacy data, the body weight used to calculate prandial, basal and total daily insulin dose per unit body weight (U/kg). If the body weight value is missing, the body weight in previous visit will be used. It means the body weight will do LOCF.

7.2.3 Handling of missing or incomplete dates

Imputation rules for missing or partial AE start dates are defined below:

If only Day of AE start date is missing:

If the AE start year and month are the same as that for the first dose date, then:

- If the full (or partial) AE end date is NOT before the first dose date or AE end date is missing, then impute the AE start day as the day of first dose date;
- Otherwise, impute the AE start day as 1.

If Day and Month of AE start date are missing:

If AE start year = first dose year, then:

- If the full (or partial) AE end date is NOT before the first dose date or AE end date is missing, then impute the AE start Month and Day as the Month and Day of first dose date
- Otherwise, impute the AE start MONTH as January and the DAY as 1.

If Year of AE start date is missing:

If the year of AE start is missing or AE start date is completely missing, then query site and leave as missing.

For missing and partial adverse event end dates, imputation will be performed as follows:

If only the day of the month is missing, the last day of the month will be used to replace the missing day. If the day and month are missing or a date is completely missing, it will be considered as missing.

For the purpose of the derivation of adverse event duration, the following rules will be applied for missing/invalid onset/resolution time:

If onset time is collected and missing/invalid, onset time will be temporarily set to 00:00:01

If resolution time is collected and missing/invalid, resolution time will be temporarily set to 23:59:59

Imputation rules for missing or partial AE stop dates and the "continuing" variable indicated as "no" are defined below:

- Year is missing – date left missing.
- Month is missing – impute December.
- Day is missing – impute last date of that month.

If imputed AE stop date is before AE start date, AE stop date will be set to AE start date.

Imputation rules for missing or partial medication start dates are defined below:

If only Day of CM start date is missing:

If the CM start year and month are the same as that for the first dose date, then:

- If the full (or partial) CM end date is NOT before the first dose date or CM end date is missing, then impute the CM start day as the day of first dose date;
- Otherwise, impute the CM start day as 1.

If Day and Month of CM start date are missing:

If CM start year = first dose year, then:

- If the full (or partial) CM end date is NOT before the first dose date or CM end date is missing, then impute the CM start Month and Day as the Month and Day of first dose date
- Otherwise, impute the CM start MONTH as January and the DAY as 1.

If Year of CM start date is missing:

If the year of CM start is missing or CM start date is completely missing, then query site and leave as missing.

If the stop date for CM is missing or partially missing and the "ongoing" variable is indicated as "no", the imputation rule is applied in the following order:

- Year is missing – Date is left missing.
- Month is missing – impute December.
- Day is missing – impute last date of that month.

7.3 Handling of site with GCP violation

During routine monitoring of this study, it was observed that one site (Site ID: 6103) had GCP compliance issues, which raised doubts on the study conduct. The blinded ADA results from the site also confirmed the need to exclude the site from results reporting. Mylan decided to conduct statistical analyses to exclude this site from ITT, Safety, and PP analysis sets. Additional sensitivity analysis to include this site (Randomized analysis set) for major baseline findings, efficacy and safety analysis will be performed. The purpose of the sensitivity analysis is to determine if major efficacy and safety analyses remain consistent with inclusion data from this site.

8 STATISTICAL METHODS

8.1 General statistical conventions

All statistical procedures will be completed using SAS version 9.3 or higher.

Unless otherwise stated, all statistical testing will be two-sided and will be performed using a significance (alpha) level of 0.05. Two-sided 95 % confidence intervals (CI) will be provided when relevant. 4 digits after decimal point will be used for all p-values.

All quantitative endpoints will be summarized using an 8-number summary (n, mean, standard deviation, median, 25th quartile, 75th quartile, minimum and maximum values). Unless otherwise specified, minimum and maximum will be presented with same number of decimal places as reported/collected, one additional decimal place for mean, median, 25th quartile and 75th quartile, and two additional decimal places for standard deviation. All qualitative endpoints will be summarized by the number of subjects meeting the endpoint and the percentage of subjects out of the appropriate population. The denominator will be displayed when needed.

For retests, unscheduled visits and visit end of treatment, in general, for by-visit summaries, data recorded at the nominal visit will be presented. Unscheduled measurements will not be included in by-visit summaries but will contribute to the scheduled visit if scheduled visit value is missing or the Last Observation Carried Forward (LOCF) value in case primary sensitivity analysis. The same principle will be applied to the visit end of treatment data.

Early termination data will be entered under the Visit 14 (week 18), but for summary tables it will be mapped to the next available planned visit number after their last scheduled visit. For example, should a patient discontinue at Visit 9, the Early Termination visit will have their Fasting plasma glucose and their HbA1c assay taken. The fasting plasma glucose would then be summarized under Visit 10 and the HbA1c under Visit 11.

All subject data, including those derived, will be presented in individual subject data listings. Unless otherwise stated, unscheduled visit results will be included in date/time chronological order, within subject listings only. All listings will be sorted by investigational site, patient number, date/time and visit. The treatment group as well as patient's sex and age will be stated on each listing. Unless otherwise stated, data listings will be based on all subjects randomized.

The visit schedules and window are shown below. If multiple records in the same visit, please use the latest one for analysis.

Study Periods	Screening	Run-in Period				Treatment Period	
Visit	V1	V2	V3	V4	V5	V6	V7
Study Week	-7 to -4	-4	-3	-2	-1	0	1
Study Day	-49 to -28	-28+/-3	-21+/-3	-14+/-3	-7+/-3	0+/-3	7+/-3
Study Day Range	[-49, -28]	[-27, -25]	[-24, -18]	[-17,-11]	[-10, -4]	[-3, 3]	[4,10]

Study Periods	Treatment Period (continue to upward table)								Follow-up
Visit	V8	V9	V10	V11	V12	V13	V14	V15	V16
Study Week	2	4	8	12	16	20	22	24	28
Study Day	14+/-3	28+/-3	56+/-7	84+/-7	112+/-7	140+/-7	154+/-7	168+/-7	196+/-7
Study Day Range	[11, 20]	[21, 40]	[41, 70]	[71,98]	[99,126]	[127,147]	[148,161]	[162, 182]	[183, 203]

8.2 Investigator pooling for secondary analyses

Investigator pooling will be done for secondary analyses where model contain investigator effect.

Pooling will be based on:

1) sites with less than 16 subject within each state will be pooled together geographically but can't over 40 subjects;

2) if a state with more than 40 subjects, sites within the state can be pooled geographically or sites will be pooled with neighbor state(s) geographically;

3) if a state with less 16 total subjects, then the sites will be pooled with other state (but can't be exceeded 40 subjects) geographically.

Pooling will be done prior to database lock and will be documented in the final BDR report.

8.3 Subject disposition

The number of patients screened, run-in, randomized, and included in each analysis population, along with study completion status, will be summarized by treatment group as well as overall. In addition, the number of patients who discontinue from the study and from IMP will be summarized by discontinuation reason. And the number of patients who discontinue from the run-in period will be summarized by discontinuation reason too.

The denominator used for the calculation of percentages will be the number of subjects randomized.

8.4 Protocol deviations

All protocol deviations identified will be summarized by treatment group and overall and by classification, i.e., major and minor.

Summaries will be conducted on all subjects that were randomized.

8.5 Demographics and baseline characteristics

No statistical testing will be performed for the comparison between treatment groups on demographics and baseline characteristics.

8.5.1 Demographics

Subject demographics will be summarized by randomized treatment group, as well as overall for all subjects in the ITT analysis population. Age, height, weight and BMI at baseline will be summarized using an 8-number summary (n, mean, standard deviation, median, 25th quartile, 75th quartile, minimum and maximum values). Qualitative variables such as gender, child-bearing potential status, ethnicity and race will be summarized using count and percentage.

A by-subject listing will be provided.

8.5.2 Baseline and disease characteristics

The categorical baseline characteristics such as duration of DM, time of basal dose, TEAR, HbA1c, FPG, insulin used prior to screening, baseline ECG and other baseline lab parameters will be summarized using frequency counts for the ITT analysis population.

A by-subject listing will be provided.

8.5.3 Medical history

A summary of medical history will be summarized by treatment group, as well as overall, by system organ class (SOC) and preferred term (PT) using latest version of Medical Dictionary for Regulatory Activities® (MedDRA) for the ITT analysis population.

A by-subject listing will be provided.

8.5.4 Concomitant medications

Medications used in this study will be coded by using the latest available version of the World Health Organization Drug Dictionary Enhanced (WHODDE).

Prior medications: are defined as medications taken within 28 days prior to screening and prior to dosing with study medication.

Concomitant medications: are defined as medications that were ongoing at the time of first dose of study medication or new medications that started after first dose of study medication and within 28 days following the date of the last dose of study medication.

Prior medications and concomitant medications will be summarized descriptively using frequency tables by anatomical therapeutic chemical (ATC) class and preferred name by treatment group, as well as overall on the ITT analysis population.

A by-subject listing will be provided.

Details for imputing missing or partial start and/or stop dates of medication are described in Section 7.2.3.

8.6 Extent of exposure

8.6.1 Treatment duration

Duration of study drug (in days) will be calculated as: last dose date – first dose date + 1 day, regardless of study drug interruption.

Study drug exposure will be summarised by treatment group as well as overall on the ITT analysis set using descriptive statistics for run-in period and treatment period.

8.6.2 Treatment compliance

Patient will be identified as study treatment non-compliant if the patient meets any following criteria:

- Missing total meal time insulin daily
- Missing basal insulin daily
- Took two times more basal insulin daily
- Took less or more than prescribed basal insulin dose units daily

Number and proportion of patients with non-compliance will be summarized along with individual categories of non-compliance for each category.

The non-compliance proportion is defined as total accumulative non-compliant days divided by duration of treatment exposure in days. Summary of non-compliance proportion will be performed by each non-compliance category for each treatment group. The treatment groups will be compared by using an ANOVA model which will include treatment and pooled-site. Treatment compliance will be summarized on the ITT analysis set.

8.7 Efficacy analyses

This section addresses separately the analyses to be conducted on the primary and secondary efficacy variables. The primary analysis population for efficacy analyses will be the ITT population, although some analyses may be repeated for other analysis sets as indicated.

8.7.1 Analysis methods

Definition of Primary Endpoints

The TEAR rate during 24-week treatment period is the primary endpoint for this study. TEAR is defined as either one of the following:

1) Subjects who are ADA negative at baseline and become positive at any timepoint post baseline.

2) Subjects who are ADA positive at baseline and demonstrate 4-fold increase in titer values at any timepoint post baseline visit.

Although differences in TEAR rate is the primary endpoint for the study from a statistical perspective, given that immunogenicity is not a standard endpoint in clinical trials, this difference, even if significant, will not be assessed in isolation but will be part of totality of evidence including changes in HbA1c, FBG, insulin dose, neutralizing antibodies and injection site reactions to ensure that changes in TEAR rate, if any, are clinically correlated and meaningful.

Statistical Methodology for Primary Endpoint

The 90% confidence interval of the treatment difference (MYL-1601D minus NovoLog®) in the TEAR rate will be established using the Wald confidence limit method with multiple imputation (Section 8.7.2.1). The MYL-1601D TEAR rate will be declared equivalent to that of NovoLog® during the 24-week treatment period if the 90% confidence interval of the treatment difference in TEAR rate is within the pre-specified margin as displayed in Table 2 of the study SAP, as suggested by Chow et al. (2002; 2003). For specific reference drug rate–based margin calculation, please refer to the SAS example code in Section 10 (Appendices – SAS Example Codes).

The primary analysis will be based on the ITT population. In addition, all post-discontinuation and rescue data will be included in the analysis.

8.7.1.1 Multiplicity

No adjustment will be made for multiple comparisons.

8.7.1.2 Treatment by center interaction analysis (multi-center study)

No analysis will be made to assess the treatment-by-center interaction.

8.7.2 Analysis of primary efficacy endpoint(s)

The primary objective of the study is to assess whether MYL-1601D is equivalent to NovoLog® in immunogenicity, as assessed by the rate of treatment emergent antibody response (TEAR) during 24 weeks of treatment. The primary efficacy endpoint is the TEAR rate at Week 24. TEAR is defined as either one of the following:

1) Subjects who are ADA negative at baseline and become positive at any timepoint post baseline
2) Subjects who are ADA positive at baseline and demonstrate 4-fold increase in titer values at any timepoint post baseline visit.

8.7.2.1 Primary efficacy analysis

The primary efficacy analysis will be a test of equivalence for the MYL-1601D and NovoLog® groups in the TEAR rate at Week 24. The null hypothesis is non-equivalence, expressed by

$$H_0\colon |p_M - p_N| \geq \delta,\ \ \delta > 0$$

where p_M and p_N are the (true) proportions of subjects with TEAR in the MYL-1601D and NovoLog® groups, respectively, and δ is the equivalence margin. The null will be rejected in favor of equivalence if the 90% confidence interval of the treatment difference is within the range $(-\delta, \delta)$. The confidence interval as defined by Chow and Shao (2002) is constructed as

$$\mathrm{CI}_{2\alpha} = [\hat{p}_M - \hat{p}_N - t_\alpha \hat{\sigma},\ \hat{p}_M - \hat{p}_N + t_\alpha \hat{\sigma}]$$

where $\hat{p}_j$ is the estimate of the TEAR rate p_j in each treatment group, $j \in \{M, N\}$, $\hat{\sigma}$ is the sample estimate of the standard error of the mean, and t_α the 0.05 quantile of the t distribution with $(N-2)$ degrees of freedom with N being the pooled number of subjects in the analysis. That is, H_0 will be rejected if neither $-\delta$ nor δ is contained within the 90% confidence interval.

To minimize any bias due to missing data, the primary efficacy analysis will use multiple imputation to fill in missing ADA values. Multiple imputation (MI) will be applied to two measured variables: ADA binary response (positive/negative) and ADA continuous titer values. Multiple imputation will be done for each variable separately, then the two sets of imputed datasets will be merged in order to generate the imputed TEAR status. For subjects who are ADA negative at baseline, post-baseline missing values will be imputed using the binary response values obtained from logistic regression MI. Missing post-baseline values for subjects who are ADA positive at baseline will be derived from the continuous MI model.

The primary efficacy analysis is based on the Week 24 (Visit 15) value. Subjects who discontinue treatment prematurely will be classified based on whether ADA values were obtained for the visits after treatment discontinuation (retrieved dropouts) or not.

For continuous-valued ADA titer, missing values prior to the last observed visit will be filled in with MCMC imputation to make the dataset monotone. Missing values after discontinuation will be imputed by regression imputation of the monotone missing dataset. If there are at least 6 retrieved dropouts in each treatment arm, then a pattern mixture model estimated from the retrieved values will be used for regression imputation of missing values. Otherwise, the complete case missing values method will be used. Missing values for the binary response variable will be imputed using a logistic regression multiple imputation model. A monotone missingness pattern will be created by taking a random draw from the binomial distribution with the same probability as the observed responses from the same treatment group at the same visit.

Pattern mixture models will be fit separately for ADA binary response and for ADA titer. Binary response will be imputed using a logistic regression imputation model, while ADA titer will use a continuous-valued regression. Independent variables in the regression will include the values at all prior visits. Each imputation will be repeated 30 times with a seed of 40872021 for ADA binary response and a seed of 41671703 for ADA titer. Code for the multiple imputation analyses is presented in Appendix 10.1.

If baseline value is missing, the missing baseline value will be imputed with six imputations based on binomial distribution using probability of no missing baseline values from same treatment group.

After imputation, the dataset containing the imputed visitwise TEAR positive or negative status binary response and the dataset containing the imputed ADA titer values will be merged by subject, visit, and imputation number. The Week 24 values will then be used to assign the status of TEAR positive or negative, as follows:

If the subject was ADA negative at baseline, then this subject will be considered TEAR positive if any post-baseline observed or imputed value was ADA positive; otherwise, the subject will be considered TEAR negative. The imputed titer values will not be considered in subjects who are ADA negative at baseline.

If the subject was ADA positive at baseline, then the imputed titer values will be used. The subject will be considered TEAR positive if the ADA titer at any post-baseline visit was ≥ 4× the baseline titer. If none of the post-baseline titer values reached 4 times the baseline titer, then the subject will be classified as TEAR negative.

For each imputation, the proportion of responders, along with standard errors, will be estimated for each treatment group, and the estimate and standard error of the risk difference will be computed. These values will be combined across imputations using PROC MIANALYZE. Subsequently, the 90% confidence interval for the risk difference will be computed using the Wald confidence limit method.

An equivalence margin will be established as a function of the TEAR rate among the Novolog subjects. Under the null hypothesis of equivalence (i.e., $H_0: p_M = p_N$), the margin can be obtained as a function of the TEAR rate in the Novolog group using the following expression (Chow S-C et al., 2003, p. 91).

$$\delta = (z_\alpha + z_\beta)\sqrt{\frac{2p_N(1-p_N)}{n_N}}$$

Selected values for δ are provided in Table 2 of Section 5 above. If the margin falls within the 90% confidence interval for the risk difference, then equivalence of MYL1601-D and Novolog will be declared.

These steps can be executed using code similar to the following:

```
proc freq data=simdata ;
title "90% Wald confidence interval";
by _imputation_;
ods output riskdiffcol2=rdiff1;
table group*resp / riskdiff (CL=WALD column=2)  alpha=0.1;
weight count;
run;

data rdiff1; set rdiff1;
keep _imputation_ row risk ase lowercl uppercl;
if row='Difference';
run;

proc mianalyze data=rdiff1;
modeleffects risk;
stderr ase;
ods output ParameterEstimates=mirisk1;
run;

ods listing;
data parms1;
set mirisk1;
lowercl = estimate - (1.645*stderr);
uppercl = estimate + (1.645*stderr) ;
run;
proc print data=parms1;
var estimate stderr lowercl uppercl;
run;
```

8.7.2.2 Secondary/Sensitivity analyses of the primary efficacy endpoint

Several sensitivity analyses will be conducted on the primary efficacy analysis:

1) The primary efficacy analysis will be repeated using observed data without imputation. This analysis will be conducted on both the ITT and the PP populations.
2) Subjects in the MYL-1601D group who are missing ADA values, and thus a TEAR assessment, will have their TEAR status at Week 24 imputed as positive. No imputation will be done for missing values in the NovoLog® group.
3) A tipping point analysis will be implemented following Liublinska and Rubin (2014). See below.

The ADA TEAR rate for each treatment group, and the difference between treatments, for

subjects who are missing data during the treatment period will be summarized for all sensitivity analyses.

The number and percent of subjects with TEAR will be summarized by treatment for each study visit during the treatment period by using similar TEAR definition: ADA negative at baseline become positive at specific visit, and ADA positive at baseline and become 4-fold increase in titer at specific visit.

A by-subject listing will be provided for observed and derived variables.

8.7.2.3 Tipping point analysis

A tipping point analysis will be used to estimate the point at which the equivalence boundary is crossed and the results fail to reject non-equivalence; that is, where

$$|\hat{p}_M - \hat{p}_N| \geq \delta, \ \ \delta > 0$$

The tipping point analysis will consider all possible counts of TEAR positive subjects that could result if TEAR status could be observed for each missing value. This will range from all missing values being imputed as positive to all missing values imputed as negative. Following Liublinska and Rubin (2014), a graphical display (heat map) is used to indicate the tipping point analysis results for each combination of imputed missing values. The risk difference in TEAR rate across the treatment arms is presented for each combination in the heat map. The graph will display the sensitivity across the range of missing values as a color gradient from most favorable for the test treatment arm to the most favorable for the reference treatment.

In addition, the TEAR rate at each scheduled visit (visitwise TEAR [vTEAR]) will be summarized at each scheduled visit by using the following criteria: Subjects who are ADA negative at baseline and become positive at a specific visit (Treatment Induced TEAR), and subjects who are ADA positive at baseline and obtain a 4-fold increase in titer at a specific visit (Treatment Boosted TEAR), will each be counted as TEAR positive at the specified visit.

8.7.3 Secondary endpoints

The following efficacy measures (both actual and change values) will be summarized at baseline and each scheduled visit.

- Change in HbA1c from baseline.
- Change in fasting plasma glucose from baseline.
- Change in prandial, basal insulin and total daily insulin dose per unit body weight (U/kg) from baseline.
- Change in 7-point self-monitored blood glucose (SMBG) profile from baseline.

A mixed model for repeated measures (MMRM) will be performed for continuous variables without imputing missing values. The MMRM model will include the fixed, categorical

effects of treatment group assignment, visit, treatment group-by-visit interaction, basal insulin dose time (AM/PM), and investigative site, as well as the baseline value as a continuous covariate.

The MMRM model results will be presented along with Least Squares (LS) means for the difference between the MYL1601-D and NovoLog® groups, standard errors, and two-sided 95% CIs for the difference between treatment groups. Contrasts of LS means at each scheduled visit will be used to evaluate all pairwise treatment comparisons, and 95% confidence intervals for treatment differences in LS means will be computed for each visit. The interaction of time points and treatment groups will also be presented. An unstructured covariance matrix will be assumed for within-subject error. The denominator degrees of freedom will be calculated according to the Kenward-Roger method.

In case of non-convergence of the above model using an unstructured covariance matrix, the following variance structures will be tested, and based on Akaike's information criteria the best fitting model will be chosen. The variance structures that will be considered for the within-subject variation in case of non-convergence are in this order: Heterogeneous Toeplitz, Toeplitz, and Compound Symmetry.

If convergence is still not obtained, the model may be revised to remove investigative site from the list of covariates.

All of the above analyses will be performed on the ITT set.

8.7.3.1 HbA1c

HbA1c is collected at week -7 to -4 (visit 1), baseline, week 12 (visit 11), and week 24 (visit 15). The measurements are performed by the central laboratory.

Actual values and changes from baseline at each time point will be computed and summarized by treatment group. Treatment difference and 95% confidence interval will be displayed along with p-values.

A sensitivity analysis for HbA1c will be performed using multiple imputation to impute missing values with non-missing subjects within the same treatment group using a pattern mixture model with the complete-case missing values (CCMV) method (Little 1993). After imputation, treatment difference, 95% CI, and p-values will be generated using same MMRM model as above. Sample imputation code is provided in Appendix 10.1.4.

8.7.3.2 Fasting plasma glucose

The fasting plasma glucose is collected at week -7 to -4 (visit 1), baseline, week 4 (visit 9), week 8 (visit 10), week 12 (visit 11), week 16 (visit 12), week 20 (visit 13), and week 24 (visit 15, End of Treatment). The measurements are performed by the central laboratory.

Actual values and changes from baseline at each time point will be computed and summarized by treatment group.

A separate analysis will be performed excluding patients with no fasting status.

8.7.3.3 Daily insulin dose per unit body weight (mealtime, basal insulin and total) for days of 7-Point profiles

Patients will be asked to document the dose of insulin taken on the same day they will perform the 7-Point SMBG profiles (from week -2 to week 24). The measurements are taken the week prior to the visits, so the week prior to:

- Week -2 (Visit 4)
- Week 0 (Visit 6)
- Week 2 (Visit 8)
- Week 4 (Visit 9)
- Week 8 (Visit 10)
- Week 12 (Visit 11)
- Week 16 (Visit 12)
- Week 20 (Visit 13)
- Week 24 (Visit 15)

Doses will be recorded in the patient's diary. This data will be transcribed to the eCRF after the patient diary is collected, by the investigator or designee at scheduled visits. The patient should document the 3 pre-visit estimations in the diary; and will be advised to document any additional estimations.

There will be three doses variables: meal time insulin dose, basal insulin and total insulin dose.

Daily Meal time insulin dose: is the sum of all the total daily meal time insulin (Humalog) dose over the double-blind treatment period divided by the number of days measured. For the computation of total daily mealtime doses, only the patients with at least three mealtime doses at a particular day are considered. That is, total daily mealtime dose will not be computed for a day if the mealtime dose is recorded less than 3 times for that day.

Daily Basal insulin dose: is the average of the insulin dose over days measured during the double-blind period.

Daily Total insulin: is the sum of all total daily meal time doses and all basal insulin doses over the double-blind treatment period divided by the number of days measured. Similarly to total daily mealtime dose for which less than 3 mealtime doses are recorded, the missing insulin doses will not be considered for the average computation. That is, if either the mealtime doses are recorded less than 3 times or the basal insulin dose is missing for a particular day of collection, total insulin dose will not be computed. The total insulin dose will be computed for the remaining days where basal insulin dose and at least 3 mealtime doses are collected.

All doses will be divided by total body weight (kg) to convert to daily insulin dose (U) per unit body weight (kg). If body weight at that visit is missing, the weight from previous visit will be used.

8.7.3.4 7-Point Self-Monitored Blood Glucose (SMBG) profile

The 7-point SMBG profile recorded via glucometer and transferred by the subject to the diary, measurement will be performed by the subject at home on 3 days (of which 2 days are consecutive) in the week before the visit

The following summaries will be summarized at week 2, 4, 8, 12, 16, 20 and week 24:

Individual 7-Points SMBG:

The over 3-day averages will be performed for the measures taken at:

- Before breakfast
- Two hours after breakfast
- Before lunch
- Two hours after lunch
- Before dinner
- Two hours after dinner
- Just before sleep

Averages will be performed at each visit.

Excursions and averages in SMBG:

The following averages will be computed:

- Morning excursion
- Noon excursion
- Evening excursion
- Overall excursion
- Pre-meal average
- Post-meal average
- (4-point average) pre-meal and bedtime average
- Overall average

The individual excursion values will be obtained by subtracting post-meal values with pre-meal values, it will be applicable for morning excursion, noon excursion and evening excursion on the respective days. For the overall excursion at morning, afternoon and evening, the individual excursion values obtained on each day will be averaged across the three days.

To compute the overall excursion, first the average of the morning, noon and evening excursion within a day will be computed, then the average will be done across the 3 days.

The pre-meal average will be the average of the following timepoints: before breakfast, before lunch, before dinner on each day, then the average will be done over the 3 days.

The post-meal average will be the average of the following timepoints: two hours after breakfast, two hours after lunch, two hours after dinner on each day, then the average will be done over the 3 days.

The pre-meal and bedtime average will be the average of the following timepoints: before breakfast, before lunch, before dinner, just before sleep, then the average will be done over the 3 days.

The overall average will be the average of all the 7 timepoints during a day. The daily average will not be computed for a day if more than 3 timepoints are missing on a particular day. The overall average will be kept missing if the daily average is missing for all three days. If the average is missing for one or two days, the average of the remaining days will be considered for the computation.

8.7.4 Analysis of exploratory endpoint(s)

Not Applicable.

8.7.5 Subgroup analyses

Subgroup analyses of important factors, including but not limited to factors such as age group, gender, race, and ethnicity are planned for the key outcomes of TEAR rate and immunogenicity variables. Wald confidence limit method for TEAR rate will be used for 90% confidence interval. Other exploratory subgroup analyses may be performed, as deemed appropriate.

The following subgroups will be assessed and described within the subgroup analysis. A minimum of 8 subjects/arm in each subgroup is needed. For subgroup with less than 8 subjects/arm, this subgroup will be pooled with the other ones having less than 8 subjects/arm together and the subgroup will be named “other” if there are other subgroups with less than 8 subjects/arm exist.

- Gender:
  - Female
  - Male
- Age (years):
  - ≤ 21
  - > 21 to <65
  - ≥65
- Race:

- American Indian or Alaska Native
- Asian
- Black or African American
- Native Hawaiian or Other Pacific Islander
- White

- Ethnicity:
  - Hispanic or Latino
  - Not Hispanic or Latino
  - Not reported or Unknown

8.8 Safety analyses

Safety data from the run-in period will be summarized and listed on run-in analysis set.

Descriptive statistics will be provided for the following safety data. No inferential analysis of this safety data is planned. Any Vital signs, ECG, and Laboratory abnormalities of potential clinical concern will be described.

For the continuous safety measures, both actual and change values will be summarized at baseline and scheduled visit. Similar MMRM model as efficacy analysis will be conducted for hypoglycemia rate. For antibody continuous variables, the MMRM model will include the fixed, categorical effect of treatment group assignment, visit, treatment group-by-visit interaction and the other fixed effect terms investigator and baseline value as covariates. Contrasts of LS mean at each scheduled visit will be used to evaluate all pairwise treatment comparisons, and 95% confidence intervals for treatment differences in LS means will be computed for each visit. For categorical data analyses, Fisher's exact test or Chi-squared test will be used.

The following analyses will also be performed.

- Incidence of ADA response, incidence of positive cross-reactive ADA, and incidence of positive NAbs
- Analysis of potential impact of TEAR status on efficacy and safety parameters such as HbA1c, FPG, insulin dose hypoglycemic rate, and incidence of any injection site allergic reactions and hypersensitivity
- Change in hypoglycemia rate (30 day adjusted) from baseline and incidence of hypoglycemic events at scheduled visits
- Incidence of TEAEs and SAEs
- Incidence of local allergic reactions, systemic allergic reactions and other adverse events
- Incidence of device-related safety assessment

To explore the potential ADA impact on subject's glucose control, the following analyses will be performed:

- The incidence of subjects with each of the three following criteria will be summarized descriptively (for categorical measures) by treatment:
    1) Meet TEAR criteria
    2) Increases in HbA1c of over 0.2% from baseline
    3) Increase in total dose from baseline
- Scatter plots of maximum tear values with HbA1c FPG and total daily insulin dose by treatment.

All the above analyses will be performed on the safety set.

In addition, graphical visualization of relationship between adverse events and treatment duration will be provided.

8.8.1 Adverse events

The frequency of AEs, SAEs, and AEs leading to discontinuation through Week 24 will be summarized based on safety analysis set or run-in analysis set.

All Adverse Events (AEs) will be classified by Primary System Organ Class (SOC) and Preferred Term (PT) according to the Medical Dictionary for Regulatory Activities (MedDRA) Version latest version. All subjects in the safety analysis set will be included in the summaries.

AEs will be classified as treatment emergent adverse events (TEAEs) are defined as follows:

TEAE: All AEs that occur after the first dose of study treatment medication after randomization through follow-up visit or 14 days after last dose [for subjects that do not have a follow-up visit] will be considered treatment emergent AEs.

Treatment-Related AEs: AE will be defined as related if causality is either definitely, probably or possible. AE will be defined as unrelated if causality is unlikely or not related. AEs where the causality is missing will be assumed to have "Reasonable possibility of relatedness."

Grade AEs:

Grade AEs (serious and non-serious) in accordance with the NCI/CTCAE scale (available at

https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_5.0/) as presented below:

- **Mild** (Grade 1) asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated
- **Moderate** (Grade 2) minimal, local or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living
- **Severe** (Grade 3) Severe or medically significant but not immediately life threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care activities of daily living

- **Life-threatening** (Grade 4) Life-threatening consequences; urgent intervention indicated
- **Death** (Grade 5) Death related to the AE.

Details for imputing missing or partial start dates of adverse events are described in Section 7.2.2. Imputed Adverse Event dates will be used for determining treatment-emergence.

Summaries of AEs will include the following:

- Treatment-emergent AEs
- Treatment-emergent treatment-related AEs
- Treatment-emergent AEs leading to study drug discontinuation
- SAEs
- Treatment-related SAEs
- AEs leading to death
- TEAE with the local and systemic allergic reactions

All TEAEs will be summarized by SOC, PT and treatment group as well as overall using frequency counts and percentages. In addition, an overall summary for the categories above will be prepared by treatment group and overall.

The number and percentage of subjects with at least one treatment emergent AE will be presented by treatment group as well as overall and events further summarized by maximum severity and relationship to study medication.

Where a subject has the same adverse event, based on preferred terminology, reported multiple times in the treatment period, the subject will only be counted once at the preferred terminology level in adverse event frequency tables.

Where a subject has multiple adverse events within the same system organ class in the treatment period, the subject will only be counted once at the system organ class level in adverse event frequency tables.

When reporting adverse events by severity, in addition to providing a summary table based on the event selection criteria detailed above, summary table will also be provided based on the most severe event during the treatment period - independent of relationship to study treatment.

Graphical visualization of relationship between adverse events and treatment duration will be provided. There are 2 graphs which are Time to AE Occurrence and Most Frequent On-Therapy AE Sorted by Risk Difference.

For run-in period, AEs and SAEs will also be summarized for run-in analysis set.

8.8.2 Device-related Assessment include both TEAE and device complaints

The total incidence of Device-related safety events will be summarized for each treatment group as well as overall and would include device-related TEAEs and events related to

device complaints or failures. For device-related TEAEs, two categories will be summarized for each treatment as well as overall: needle-related TEAEs such as pain, bruise, and bleeding; and other device-related TEAEs, such as hyperglycemia or hypoglycemia. For device-related patient's complaints, incidence will be listed and summarized for each treatment and overall. Observed values and change from baseline will also be presented.

All the above analyses will be performed on the safety set for treatment period and run-in period.

8.8.3 Hypoglycemia

Hypoglycemia is a state produced by a lower than normal level of glucose in the blood. Hypoglycemia is classified as severe, documented symptomatic, asymptomatic, probable symptomatic, relative, nocturnal hypoglycemia. Patients will be instructed to record all hypoglycemic events in the patient's diary from Visit 2 until the EOT visit. The hypoglycemic events will be reviewed by the investigator and transcribed into the eCRF by the investigator or designee after the diary has been collected.

The classification will be derived as follows:

- Severe: the patients entered in the eCRF: "severe (external assistance required to resolve event)"
- Documented Symptomatic Hypoglycemia: the patient entered in the eCRF: symptomatic and checked the "glucose value was less than or equal to 70 mg/dL".
- Asymptomatic Hypoglycemia: the patient entered in the eCRF: asymptomatic (symptoms of hypoglycemia not present) and checked the "glucose value was less than or equal to 70 mg/dL".
- Probable Symptomatic Hypoglycemia: the patient entered in the eCRF: symptomatic and checked the "glucose was not measured".
- Relative Hypoglycemia: the patient entered in the eCRF: symptomatic and checked the "glucose value was over 70 mg/dL".
- Nocturnal Hypoglycemia: the patient has any of the 5 types above and also checked the "Nocturnal" time of the event.

Hypoglycemia event rate per patient per 30 days calculated between two visits is defined as total number of episodes between two visits divided by the number of days between the visits, multiplied by 30 days. The baseline hypoglycemia period is defined from the run-in period until randomization day. This rate will also be calculated per patient for nocturnal hypoglycemia episodes.

Hypoglycemia event rate per patient per 30 days will be analyzed using the same MMRM model as for the primary efficacy parameter (without imputing any missing data).

For change from baseline of hypoglycemia rate, a graphical display at scheduled visits, of LSMeans and 2-sided 95% CI will be generated from the MMRM model. Additionally, mean (+/- SD) for actual measurements by visit will also be presented.

In addition, nocturnal hypoglycemia rate and incidence will be analyzed in a same way as overall hypoglycemic episodes.

Listings of hypoglycemic episodes and severe hypoglycemic episodes will be presented by visit for each patient. If a sufficient number of severe hypoglycemic episodes are reported, then incidence summaries similar to the incidence of hypoglycemic episodes will be included.

8.8.4 Hypersensitivity and Immune mediated adverse events

Incidence of hypersensitivity and immune mediated adverse events will be presented for treatment groups as well as overall.

8.8.5 Immunogenicity assessments

For antibody continuous variables, the MMRM model will include the fixed, categorical effect of treatment group assignment, visit, treatment group-by-visit interaction and the other fixed effect terms investigator and baseline value as covariates. Contrasts of LS mean at each scheduled visit will be used to evaluate all pairwise treatment comparisons, and 95% confidence intervals for treatment differences in LS means will be computed for each visit. For categorical data analyses, Fisher's exact test or Chi-squared test will be used.

The following analyses will also be performed.

- Incidence of ADA by visit, incidence of positive cross-reactive ADA by ADA status and by visit, and incidence of NAbs.
- Analysis of potential impact of TEAR status on efficacy and safety parameters such as HbA1c, glucose control FPG, insulin dose, hypoglycemic rate, and Incidence of any injection site allergic reactions and hypersensitivity.

To explore the potential ADA impact on the subjects' glucose control, scatter plots of maximum TEAR values with HbA1c, FPG and total daily insulin dose will be displayed by treatment. In addition, the following analyses will be performed:

- TEAR effect on Efficacy will be assessed using the following:
  1) HbA1c and total insulin dose values by TEAR status and treatment group.
  2) A summary of subjects who experience meet vTEAR criteria, an increase in A1c of at least 0.2% from baseline, and an increase in the total dose from baseline at a protocol scheduled visit (where immunogenicity, A1c and dose all measured). This summary will be provided by visit, by treatment group, and overall.
- TEAR effect on Safety will be assessed using the following:

1) Hypoglycemic rate by treatment group and TEAR status.
2) Incidence of injection site allergic reactions and incidence of hypersensitivity in TEAR positive subjects.
3) Incidence of positive NAb by TEAR status and treatment group.

All the above analyses will be performed on the safety set.

The NAb actual assay results will not be available at the time of database lock (EDC database lock) for the analysis. However, the treatment will remain blinded during the laboratory assay. These data will be sent later from assay lab and statistical analysis will be performed according to pre-specified method in this SAP.

8.9 Other Safety Analyses

8.9.1 Laboratory Data

Change from baseline of laboratory measurements will be analyzed using MMRM with model terms of investigator, treatment, visit, treatment-by-visit interaction as fixed effects, and baseline value as covariate. The descriptive statistics for actual measurement and change from baseline along with treatment comparison will be performed at scheduled visits.

The percentage of subjects in categories such as normal, abnormal/non-clinically significant and abnormal/clinically significant will be summarized by treatment.

The percentage of subjects who meet potentially clinically significant criteria will be summarized by treatment groups.

The criteria for the Potential Clinically Significant Lab is listed in Appendix 10.3.

The following safety laboratory tests will be performed at times defined in the study schedules in Table 1.

Table 4: Laboratory Safety Tests

Hematology	Chemistry	Urinalysis	Other
Hemoglobin Hematocrit RBC count Platelet count WBC count Total neutrophils (Abs) Eosinophils (Abs) Monocytes (Abs) Basophils (Abs) Lymphocytes (Abs)	Urea and Creatinine Glucose Calcium Sodium Potassium Chloride AST, ALT Total Bilirubin Direct/Indirect bilirubin Alkaline phosphatase Uric acid Albumin Total protein CRP C-Peptide Lipid Profile	pH Glucose (qual) Protein (qual) Blood (qual) Ketones Nitrites Leukocyte esterase Microscopy/culture[a]	Urine/Serum hCG HIV and HBsAg and HCVAb HbA1c
[a] Only if urine dipstick is positive for blood, protein, nitrites or leukocyte esterase.			

Hematology, chemistry, urinalysis and other (parameters in last column of Table 4) will be summarized for treatment groups as well as overall by visits and parameters in each category.

8.9.2 Vital signs

Change from baseline of vital sign measurements will be analyzed using MMRM with model terms of investigator, treatment, visit, treatment-by-interaction as fixed effects, and baseline value as covariate. The descriptive statistics including actual measurement and change from baseline along with treatment comparison will be performed at scheduled visits.

The percentage of subjects who meet potentially clinically significant criteria will be summarized by treatment groups.

8.9.3 Physical examinations

A full physical examination will be performed at Visit 1 (screening), Visit 6 (randomization) and Visit 15 (EOT or ET). Height and weight will be assessed at Visit 1. The percentage of subjects in categories such as normal, abnormal/non-clinically significant and abnormal/clinically significant will be summarized at scheduled visits. But clinically significant changes from the screening procedures results will be recorded as adverse events. Physical examinations will be summarized for the safety analysis set.

8.9.4 Electrocardiograms (ECG)

The percentage of subjects in categories such as normal, abnormal/non-clinically significant and abnormal/clinically significant will be summarized by treatment group at screening and EOT.

8.10 Interim analysis

No interim analysis is planned for this study.

9 REFERENCES

1) ICH Topic E3: Structure and Content of Clinical Study Reports (CPMP/ICH/137/95-adopted December 1995).

2) Dong and Peng. Principled missing data methods for researchers. SpringerPlus, 2013, 2:222.

3) Chow S-C and Shao J. A note on statistical methods for assessing therapeutic equivalence. Controlled Clinical Trials, 2002; 23: 515–520.

4) Chow S-C, Shao J, Wang H. Sample Size Calculations in Clinical Research. New York: Marcel Dekker, 2003.

5) Liublinska and Rubin. Sensitivity analysis for a partially missing binary outcome in a two-arm randomized clinical trial. Stat Med. 2014; 33(24) 4170-4185.

6) Roderick J. A. Little. Pattern-Mixture Models for Multivariate Incomplete Data. Journal of the American Statistical Association. 1993: 88: 125-134.

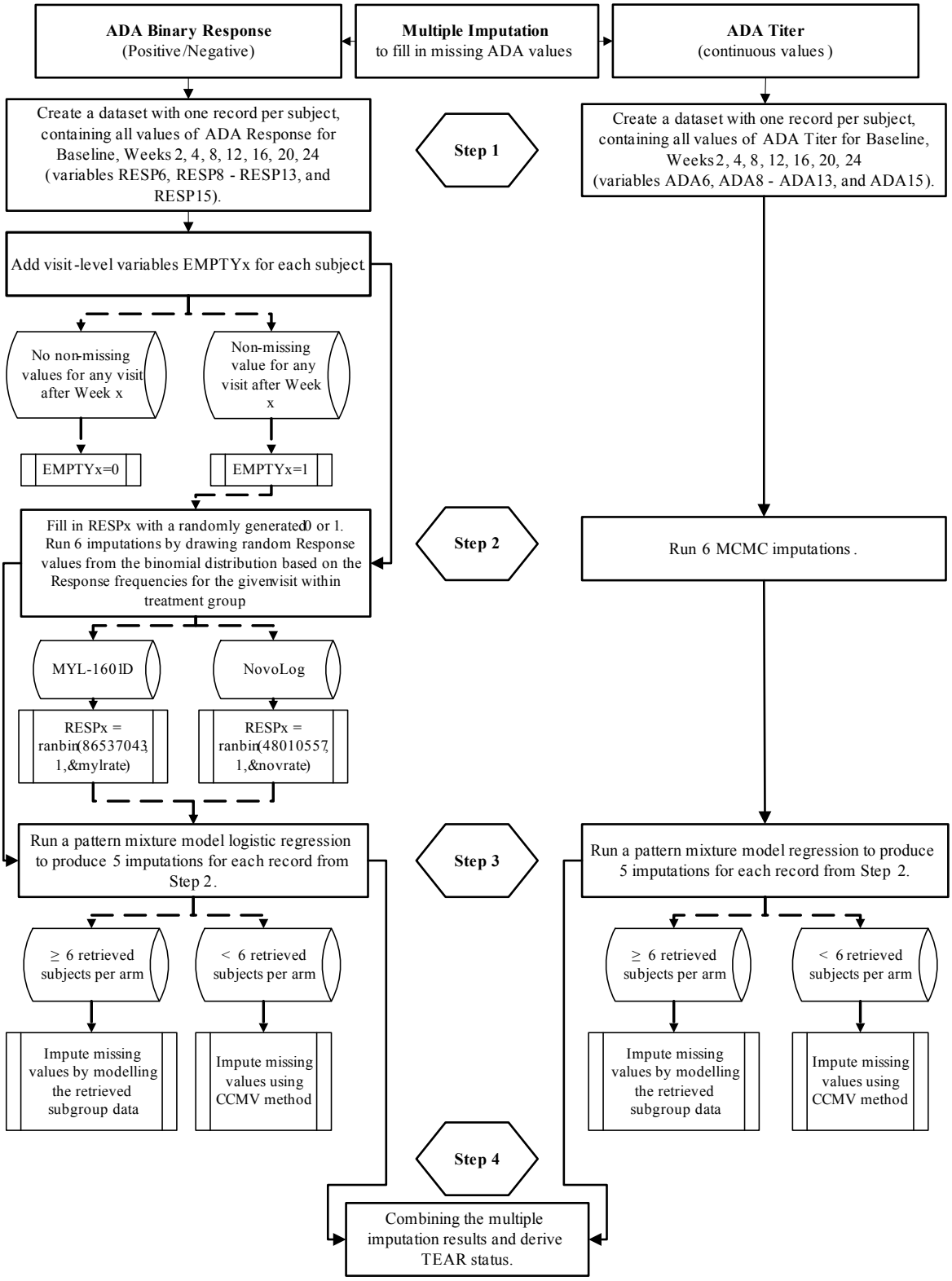
10 APPENDICES – SAS EXAMPLE CODES

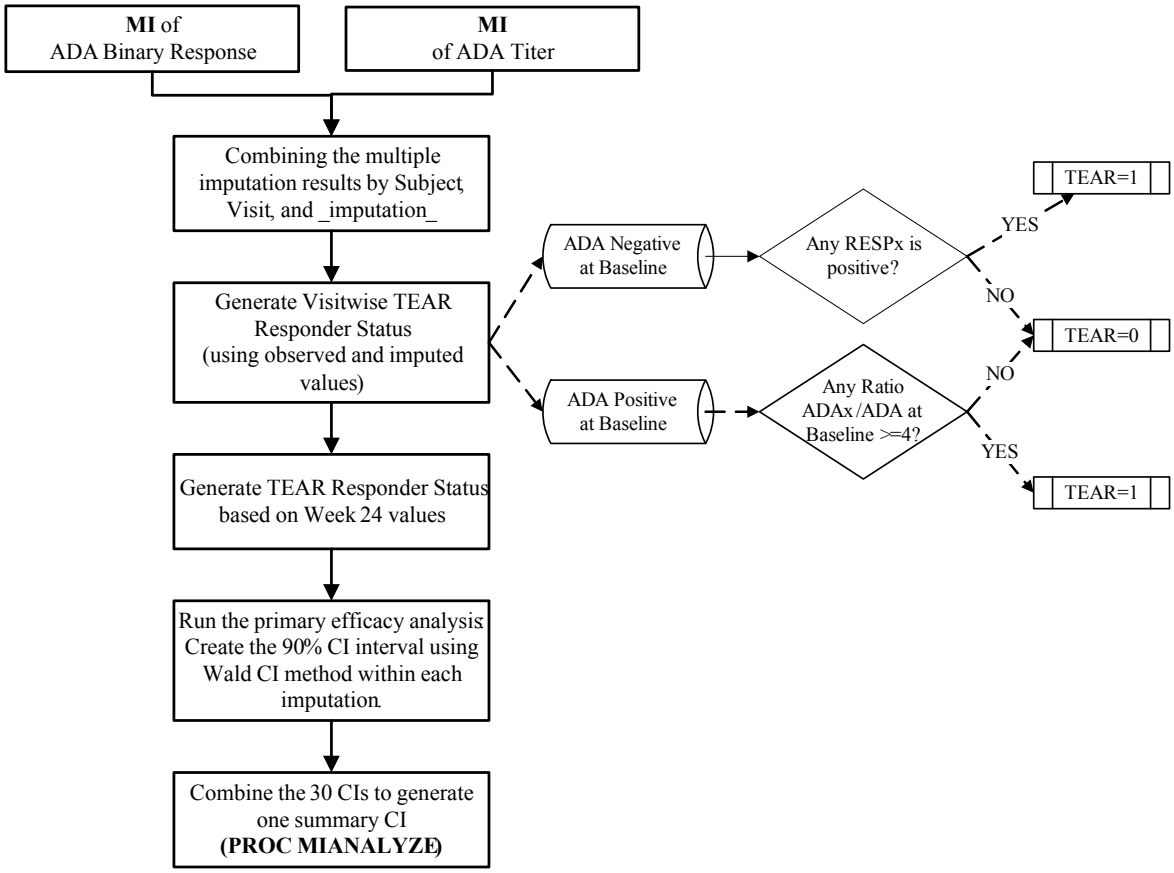
10.1 Multiple imputation

Multiple imputation will be applied to fill in missing values for ADA binary response and for ADA titer.

For binary ADA response, if baseline ADA value is missing, the missing baseline value will be imputed (6 imputations) based on binomial distribution using probability of no missing baseline values form the same treatment group. After imputations, for subjects with negative baseline ADA (Section 10.1.1) missing post-baseline binary response (positive or negative) will be imputed using logistic regression multiple imputation. For subjects who are ADA positive at baseline (Section 10.1.2), missing post-baseline ADA continuous titer values will be imputed with same treatment group non-missing subjects using continuous MI model.

The steps are summarised below:


ADA Binary Response
(Positive/Negative)
Multiple Imputation
to fill in missing ADA values
ADA Titer
(continuous values)
Create a dataset with one record per subject, containing all values of ADA Response for Baseline, Weeks 2, 4, 8, 12, 16, 20, 24 (variables RESP6, RESP8 - RESP13, and RESP15).
Step 1
Create a dataset with one record per subject, containing all values of ADA Titer for Baseline, Weeks 2, 4, 8, 12, 16, 20, 24 (variables ADA6, ADA8 - ADA13, and ADA15).
Add visit-level variables EMPTYx for each subject
No non-missing values for any visit after Week x
Non-missing value for any visit after Week x
EMPTYx=0
EMPTYx=1
Fill in RESPx with a randomly generated 0 or 1. Run 6 imputations by drawing random Response values from the binomial distribution based on the Response frequencies for the given visit within treatment group
Step 2
Run 6 MCMC imputations.
MYL-1601D
NovoLog
RESPx = ranbin(86537043, 1,&mylrate)
RESPx = ranbin(48010557, 1,&novrate)
Run a pattern mixture model logistic regression to produce 5 imputations for each record from Step 2.
Step 3
Run a pattern mixture model regression to produce 5 imputations for each record from Step 2.
≥ 6 retrieved subjects per arm
< 6 retrieved subjects per arm
≥ 6 retrieved subjects per arm
< 6 retrieved subjects per arm
Impute missing values by modelling the retrieved subgroup data
Impute missing values using CCMV method
Impute missing values by modelling the retrieved subgroup data
Impute missing values using CCMV method
Step 4
Combining the multiple imputation results and derive TEAR status.



Section 10.1.1 describes the methodology for imputation of binary response, while Section 10.1.2 covers the imputation of continuous ADA titer, and Section 10.1.3 puts them together.

10.1.1 Multiple imputation for binary ADA response if baseline ADA is negative.

For purposes of these specifications, let RESPx represent the binary value (0 or 1) at Visit x, where x takes the values 0 (for Baseline) or 8–15. The seed that will be used in the multiple imputation of ADA binary response will be 40872021.

Missing values will be obtained from logistic regression imputation on the observed values of binary response. A fully conditional specification (FCS) method will be used to impute values from a non-monotone missing data pattern.

There are 3 steps to this process:

1. Create a dataset with one record per subject, containing all values of RESP for all weeks (Baseline and Weeks 2–24 on one record); that is, each record would contain variables for RESP0 and RESP8–RESP15. The subject-level RETRIEVED variable described in Section 8.7.2.1 can also be generated at this time for subjects who discontinued treatment, based on whether values after treatment discontinuation

were retrieved (set retrieved=1) or remain missing (set retrieved=0).

2. Fill in missing values to make a monotone missing data pattern, by drawing random TEAR values from the binomial distribution based on the TEAR frequencies for the given visit within treatment group.
3. Run a pattern mixture model logistic regression to impute missing values to generate 30 imputed datasets. This will generate a single pooled dataset containing one observation per subject per imputation. All missing values will have been filled in for each subject.

Step 1.

The dataset for this imputation has to put the values for all weeks onto one record. That is, the dataset would have one record per subject, with 8 values for RESP0 and RESP8–RESP15. One way this can be accomplished is the following:

```
data analysis (keep=usubjid RESP0 RESP8-RESP14);
set indata;
by usubjid avisitn;
array R {9} RESP0 RESP8-RESP15;
retain RESP0 RESP8-RESP15;
if first.usubjid then do;
   resp0 = base;
   do visit = 8 to 15;
      R{visit} = .;
   end;
end;
do visit = 8 to 15;
   if avisit = visit then R{visit} = aval;
end;
if last.usubjid then output;
run;
```

STEP 2.

First a flag needs to be created to indicate which missing values need to be filled in. This is a visit-level variable for each subject. For explanatory purposes, this variable will be referred to as `EMPTYx` for Visit x. For each visit where a subject has a missing value, use the logic below:

- If there is a non-missing value for any visit AFTER Visit x, then `EMPTYx=0;`
- Else if there are no non-missing values for any of the remaining visits, then `EMPTYx=1;`

Then all records with EMPTYx=1 have to have their RESPx value filled in with a randomly generated 0 or 1. One way this can be accomplished is the following:

```
/*** step 2: Make missingness monotone ***/
%global mylrate novrate;

%macro getrate(visnum);

   proc sort data=indata;
   by trtgroup;
   run;
   ods exclude all;
   proc freq data=indata;
   by trtgroup;
   table RESP&visnum;
   ods output onewayfreqs=pct;
   run;
   ods exclude none;

   data binom;
   set pct (where=(RESP&visnum=1));
   probresp = percent/100;
   if trtgroup="MYL-1601D" then call
   symput("mylrate",probresp);
   else if trtgroup="NOVOLOG" then call
   symput("novrate",probresp);
   run;

%mend getrate;


%macro fillin (visit);

   %getrate(&visit);

   data indata;
   set indata;
   if empty&visit=1 then do;
      if trtgroup="MYL-1601D" then
        RESP&visit=ranbin(86537043,1,&mylrate);
      else if trtgroup="NOVOLOG" then
        RESP&visit=ranbin(48010557,1,&novrate);
```

```
   end;
   run;

%mend fillin;


%macro monotone;
   %do vis=8 %to 13;
      %fillin(&vis);
   %end;
%mend monotone;

%monotone;
```

STEP 3.

Since SAS will not accept a binary variable as an independent variable in a logistic imputation model, continuous versions of the binary RESP variables. Both the continuous and categorical variables will be used during the regression. In the code below, the logistic regression code is different If there are <6 retrieved dropouts per group versus ≥6.

```
/*** setup: create numeric versions of binary variables
***/
data indata;
set indata;
array resps {7) resp0 resp8-resp14;
array respns {7) respn0 respn8-respn14;
do i=1 to 7;
   respns{i} = resps{i};
end;
run;
```

```
/*** step 3: Perform logistic regression imputation ***/

/*** Code to use if >=6 retrieved subjects per arm ***/
proc mi data=indata nimpute=30 seed=40872021
out=outdata;
class resp8-resp15;
var resp0n resp8 resp8n resp9 resp9n resp10 resp10n
resp11
   resp11n resp12 resp12n resp13 resp13n resp14 resp14n
resp15;
monotone logistic (resp8= resp0n) ;
monotone logistic (resp9= resp0n resp8n) ;
monotone logistic (resp10= resp0n resp8n resp9n) ;
monotone logistic (resp11= resp0n resp8n resp9n resp10n)
;
monotone logistic (resp12= resp0n resp8n resp9n resp10n
resp11n) ;
monotone logistic (resp13=
   resp0n resp8n resp9n resp10n resp11n resp12n) ;
monotone logistic (resp14=
   resp0n resp8n resp9n resp10n resp11n resp12n resp13n)
;
monotone logistic (resp15=
   resp0n resp8n resp9n resp10n resp11n resp12n resp13n
resp14n) ;
mnar model (resp8-resp15 / modelobs=(retrieved=1));
run;
```

```
/*** Code to use if <6 retrieved subjects per arm ***/
proc mi data=indata nimpute=30 seed=40872021
out=outdata;
by trtgroup;
class resp8-resp15;
var resp0n resp8 resp8n resp9 resp9n resp10 resp10n
resp11
  resp11n resp12 resp12n resp13 resp13n resp14 resp14n
resp15;
monotone logistic (resp8= resp0n) ;
monotone logistic (resp9= resp0n resp8n) ;
monotone logistic (resp10= resp0n resp8n resp9n) ;
monotone logistic (resp11= resp0n resp8n resp9n resp10n)
;
monotone logistic (resp12= resp0n resp8n resp9n resp10n
resp11n) ;
monotone logistic (resp13=
  resp0n resp8n resp9n resp10n resp11n resp12n) ;
monotone logistic (resp14=
  resp0n resp8n resp9n resp10n resp11n resp12n resp13n)
;
monotone logistic (resp15=
  resp0n resp8n resp9n resp10n resp11n resp12n resp13n
resp14n) ;
mnar model (resp8-resp15 / modelobs=ccmv(k=1));
run;
```

10.1.2 Multiple imputation for continuous ADA titer if baseline ADA positive.

For purposes of these specifications, let ADAx represent the continuous value for ADA titer at Visit x, where x takes the values 0 (for Baseline) or 8–15. The seed that will be used in the multiple imputation of ADA titer will be 41671703.

There are 3 steps to this process:

1. Create a dataset with one record per subject, containing all values of ADA for all weeks (Baseline and Weeks 2–24 on one record); that is, each record would contain variables for ADA0 and ADA8–ADA15. The subject-level RETRIEVED variable described in Section 8.7.2.1 can also be generated at this time for subjects who discontinued treatment, based on whether values after treatment discontinuation were retrieved (set retrieved=1) or remain missing (set retrieved=0).
2. Run 30 Markov Chain Monte Carlo (MCMC) imputations to make the missing

values in the dataset have a monotone missingness pattern. This will produce a dataset with an index variable _imputation_ and otherwise the same variables as the input dataset. It will have one record per imputation per subject.

3. Run a pattern mixture model regression imputation on each imputed dataset from Step 2. This will generate a dataset with the same structure as the input dataset, but will have the missing values filled in.

Step 1.

The dataset for this imputation has to put the values for all weeks onto one record. That is, the dataset would have one record per subject, with 8 values for ADA0 and ADA8–ADA15. One way this can be accomplished is the following:

```
data analysis (keep=usubjid trtp ADA0 ADA8-ADA15);
set indata;
by usubjid avisitn;
array y {9} ADA0 ADA8-ADA15;
retain ADA0 ADA8-ADA15;
if first.usubjid then do;
   ada0 = base;
   do visit = 8 to 15;
      y{visit} = .;
   end;
end;
do visit = 8 to 15;
   if avisit = visit then y{visit} = aval;
end;
if last.usubjid then output;
run;
```

Step 2.

```
/*** step 2: Create monotone missingness using MCMC
imputation ***/
proc mi data=analysis seed=41671703 nimpute=1
out=outdata1;
mcmc impute=monotone;
var ada0 ada8-ada15;
run;
```

Step 3.

```
/*** step 3: Perform PMM regression imputation ***/
proc sort data=outdata1;
run;

/*** Code to use if >=6 retrieved subjects per arm ***/
proc mi data=outdata1 seed=41671703 nimpute=30
out=outdata2;
class retrieved;
monotone reg(ada8-ada15 / details);
mnar model (ada8-ada15 / modelobs=(retrieved=1));
var ada0 ada8-ada15;
run;

/*** Code to use if <6 retrieved subjects per arm ***/
proc mi data=outdata1 seed=41671703 nimpute=30
out=outdata2;
by treatment;
monotone reg(ada8-ada15 / details);
mnar model (ada8-ada15 / modelobs=ccmv(k=1));
var ada0 ada8-ada15;
run;
```

10.1.3 Combining the multiple imputation results to derive TEAR status

At this point there should be two sets of imputed data, each containing 30 records per subject, with all visits on the same record. These will need to be merged by subject and imputation. Then the TEAR values will be computed for each subject and analyzed. There are 3 steps:

1. Merge the imputed ADA binary response dataset with the imputed ADA titer dataset, by `usubjid avisitn _imputation_`. Then generate TEAR responder status based on the Week 24 (Visit 15) ADA values using the same algorithm as was used for the original data. This dataset will have one record per imputation per subject, but the TEAR values will be non-missing for all subject visits. Note that the variable TEAR may vary from imputation to imputation for the same subject. That is supposed to happen.
2. Run the primary efficacy analysis to create the 90% confidence interval using the Wald confidence limit method within each imputation. That will result in 30 confidence intervals, one for each imputation.
3. Combine the 30 confidence intervals to generate one summary confidence interval to be presented in the table.

Step 1.

Merge the imputed titer values by `usubjid avisitn _imputation_` to make a single dataset. Then create a new variable for TEAR status by following the ADaM specs for the definition of TEAR events, but using the new imputed ADA data. Again, the dataset will contain one record per imputation per subject, with a new response variable added to every record. For the purposes of the next steps below, this new response variable will be called TEAR.

a. If the subject was ADA negative at baseline then do:
   i. If any of RESP8–RESP15 is positive, then TEAR=1 for that imputation for that subject
   ii. Else TEAR=0 for that imputation for that subject
b. Else if subject was ADA positive at baseline then do:
   i. If ADAx / ADA0 >=4 for any visit, where x is in 8–15, then TEAR = 1;
   ii. else TEAR=0

Step 2.

Use code similar to what was used to analyze the primary efficacy endpoint, but run the code by `_imputation_`. For example,

```
/*** step 2: Analyze the imputed datasets ***/
proc freq data=imputed;
by _imputation_;
table trt*tear / alpha=0.10 riskdiff (cl= wald column=2)
;
output=outmi riskdiff;
run;
```

Step 3.

Combine the 30 estimated statistics into one using PROC MIANALYZE.

```
proc mianalyze data=outmi;
modeleffects _rdif1_ ;
stderr e_rdif1;
ods output parameterestimates=parms;
run;


proc print data=parms;
var estimate stderr lclmean uclmean probt;
```

```
run;
```

10.1.4 Multiple imputation for HbA1c.

HbA1c is measured at baseline and at Weeks 12 and 24 (visits 11 and 15). For purposes of these specifications, let HBx represent the continuous value for HbA1c value at Visit x, where x takes the values 0 (for Baseline) or 11 or 15. The seed that will be used in the multiple imputation of ADA titer will be 24637883. The imputations will be combined using a mixed model for repeated measures.

There are 6 steps to this process:

1. Create a dataset with one record per subject, containing all values of HbA1c for all weeks (Baseline and Weeks 12 and 24 on one record); that is, each record would contain variables for HB0, HB11, and HB15.
2. Run 30 Markov Chain Monte Carlo (MCMC) imputations to make the missing values in the dataset have a monotone missingness pattern. This will produce a dataset with an index variable `_imputation_` and otherwise the same variables as the input dataset. It will have one record per imputation per subject.
3. Run a pattern mixture model regression imputation on each imputed dataset from Step 2. This will generate a dataset with the same structure as the input dataset, but will have the missing values filled in.
4. Transform the dataset back into one record per subject per visit, with imputation number retained on each record.
5. Run a mixed model for repeated measures on the imputed datasets.
6. Combine the imputations to provide summary statistics.

Step 1.

The dataset for this imputation has to put the values for all weeks onto one record. That is, the dataset would have one record per subject, with 3 values for HB0, HB11, and HB15. One way this can be accomplished is the following:

```
data analysis (keep=usubjid trtp HB0 HB11 HB15);
set indata;
by usubjid avisitn;
array y {3} HB0 HB11 HB15;
retain HB0 HB11 HB15;
if first.usubjid then do;
   hb0 = base;
   hb11 = .;
   hb15 = .;
end;
```

```
if avisit = 11 then hb11 = aval;
if avisit = 15 then hb15 = aval;
if last.usubjid then output;
run;
```

Step 2.

```
/*** step 2: Create monotone missingness using MCMC
imputation ***/
proc mi data=analysis seed=24637883 nimpute=1
out=outdata1;
mcmc impute=monotone;
var hb0 hb11 hb15;
run;
```

Step 3.

```
/*** step 3: Perform PMM regression imputation ***/
proc sort data=outdata1;
run;

proc mi data=outdata1 seed=24637883 nimpute=30
out=outdata2;
by treatment;
monotone reg(hb11 hb15 / details);
mnar model (hb11 hb15 / modelobs=ccmv(k=1));
var hb0 hb11 hb15;
run;
```

Step 4.

Reverse the process from Step 1, creating one record per subject visit, retaining the imputation number on each record.

Step 5.

Use code similar to what was used to analyze the primary efficacy endpoint, but run the code by `_imputation_`. For example,

```
/*** step 5: Analyze the imputed datasets ***/
proc sort data=imputed;
by _imputation_ subjid avisitn;
run;

proc mixed data=imputed;
by _imputation_;
```

```
class trt avisitn dosetime site;
model chg = trt avisitn dosetime site trt*avisitn base /
ddfm=kr s;
repeated avisitn / type=un;
lsmeans trt*avisitn / diff cl;
ods output ConvergenceStatus=converge;
ods output diffs=trtdiff;
run;
```

Convergence status is assessed by the condition that:

```
if pdg=1 and pdh=1 and status=0
```

in the dataset `converge` above, then model has converged. In the above code, if the model does not converge, replace the `type=un` option by each of the following: `type=toeph`, `type=toep`, and `type=cs`. You will also need to replace `ddfm=kr` with `ddfm=bw`. Choose the best fit model by checking the AIC and selecting the model with the lowest one. Please contact the statistician if this step is necessary.

Step 6.

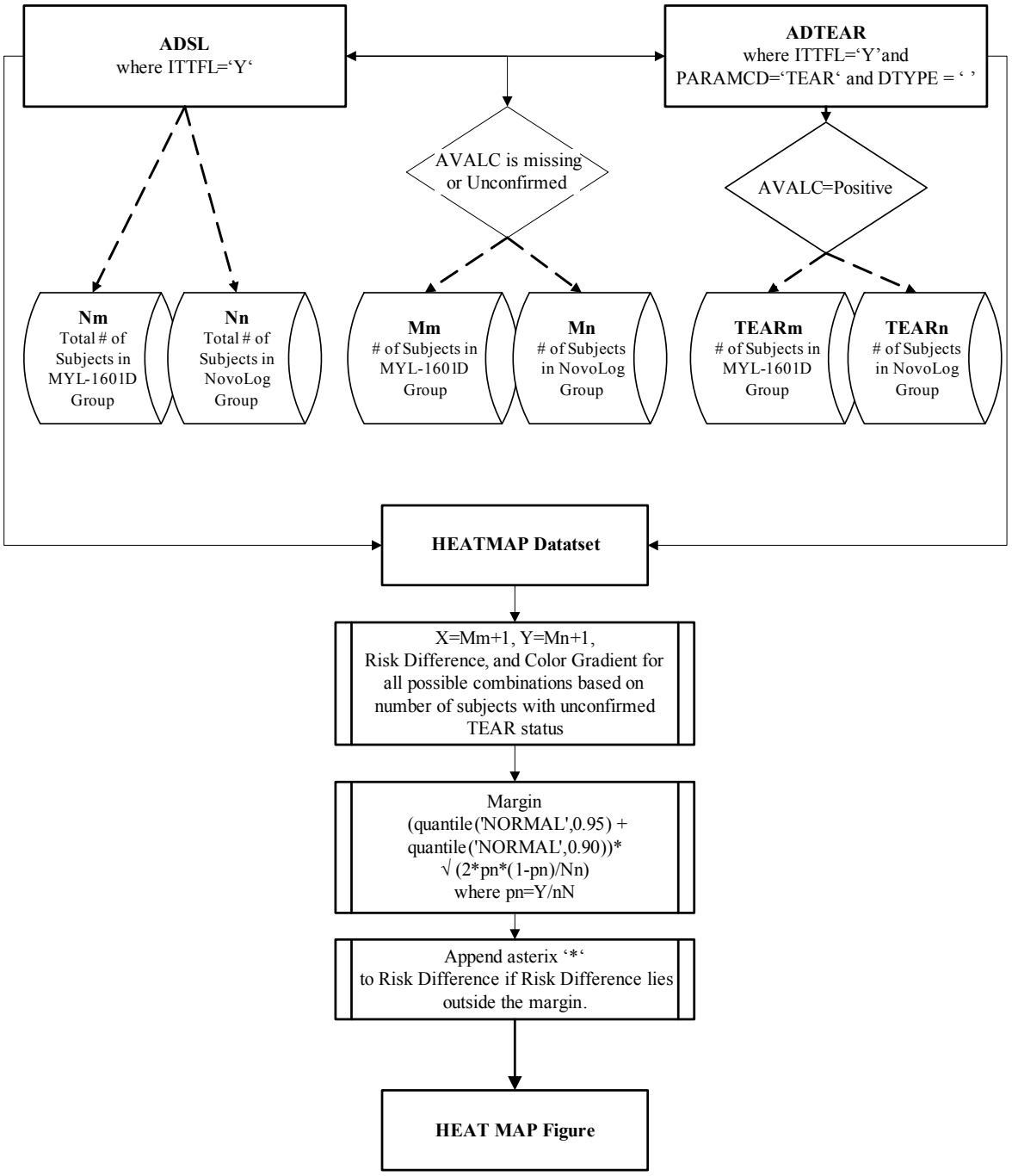
Combine the 30 estimated statistics into one using PROC MIANALYZE.

```
data trtdiff;
set trtdiff (where=(avisitn=_avisitn));
run;

proc sort data=trtdiff;
by avisitn _imputation_;
run;

proc mianalyze data=trtdiff;
by avisitn;
modeleffects estimate ;
stderr stderr;
ods output parameterestimates=parms;
run;


proc print data=parms;
var avisitn estimate stderr lclmean uclmean probt;
run;
```

10.2 Tipping Point Analysis

The tipping point analysis will create a type of figure called a heat map. It will not impute data at the subject level, but rather will be based on the observed total counts of TEAR status and number of missing values per group.

The steps are summarised below:



For purposes of the code below, set

```
Nm = Total # of subjects randomized to MYL-1601D group
Nn = Total # of subjects randomized to Novolog group
Mm = # of subjects missing TEAR status in MYL-1601D
group
Mn = # of subjects missing TEAR status in Novolog group
TEARm = # of subjects in MYL-1601D group who are TEAR
positive
TEARn = # of subjects in Novolog group who are TEAR
positive
```

Then run code similar to the following. The code below assumes that the above variables have been created as global variables.

```
%macro heat;

data heatmap;
do i = 0 to &mn;
   do j = 0 to &mm;
      y = &tearn + i;
      x = &tearm + j;
     pn = y/&nn;
     pm = x/&nm;
     riskdiff = pm - pn;
    margin = (quantile('NORMAL',0.95) +
    quantile('NORMAL',0.90))*
       sqrt(2*pn*(1-pn)/&nn);
     output;
   end;
end;
run;

%do i = 0 %to &mn;
   %do j = 0 %to &mm;

   data imputed;
   length group $8;
   group="NOVOLOG";
   status=1;
   count=%eval(&tearn+&i);
   output;
   status=0;
   count=%eval(&nn - (&tearn+&i));
   output;
```

```
    group="MYL-1601D";
    status=1;
    count=%eval(&tearm+&j);
    output;
    status=0;
    count=%eval(&nm - (&tearm+&j));
    output;
    run;

    ods exclude all;
    proc freq data=imputed;
    table group * status / riskdiff (cl=WALD column=2)
alpha=0.1;
    weight count;
    ods output riskdiffcol2=rdiff1;
    run;
    ods exclude none;
    data ci;
    set rdiff1 (where=(row="Difference"));
       y = &tearn + &i;
       x = &tearm + &j;
    run;

    data heatmap;
    merge heatmap
       ci (keep=x y risk ase lowercl uppercl)
    ;
    by x y;
   length color $8 digits $16 a1-a6 asterisk $1;
   if -0.0000001 <riskdiff < 0.0000001 then do;
      red=255;
      green=255;
      blue=255;
   end;
   else if riskdiff > 0 then do;
      if riskdiff >= 0.16 then do;
         red = round(484 - (1435*riskdiff));
     green = 0;
      blue = 0;
      end;
      else do;
         red = 254;
         green = round(255/(1+exp(-32.96*(0.1149-
abs(riskdiff)))));
         blue = round(255/(1+exp(-21.58*(0.0793-
abs(riskdiff)))));
```

```
        end;
      else if . < riskdiff < 0 then do;
         red = round(255/ (1+exp(-18.51*(0.1108-
      abs(riskdiff)))));
         green = round(255/ (1+exp(-27.85*(0.1699-
      abs(riskdiff)))));
         blue = round(255/ (1+exp(-20.10*(0.1024-
      abs(riskdiff)))));
      end;

       digits = "0123456789ABCDEF";
       n1 = floor(red/16);
       n2 = red - (n1*16);
       a1 = substr(digits,n1+1,1);
       a2 = substr(digits,n2+1,1);
       n3 = floor(green/16);
       n4 = green - (n3*16);
       a3 = substr(digits,n3+1,1);
       a4 = substr(digits,n4+1,1);
       n5 = floor(blue/16);
       n6 = blue - (n5*16);
       a5 = substr(digits,n5+1,1);
       a6 = substr(digits,n6+1,1);
       color = compress("CX"||a1||a2||a3||a4||a5||a6);

      if (lowercl < (margin) < uppercl) and
         (lowercl < (-margin) < uppercl) then asterisk=" ";
      else asterisk="*";

       run;

       %end;
    %end;
%mend heat;

%heat;
```

This code will create a dataset called HEATMAP that contains the X and Y variables, risk difference, and color to use for the plot. There will be (Mm+1) values on the x-axis and (Mn+1) values on the y-axis. These will be the number of subjects with TEAR=1 in the MYL-1601D (x-axis) and Novolog (y-axis) groups, respectively, based on all possible counts that could be observed if the missing TEAR values were known. The HEATMAP dataset provides these counts as variables x and y for each row.

The heat map figure will consist of a grid with dimensions (Mm+1) × (Mn+1). The x-axis

will be labeled with the range of values of the variable x in the HEATMAP dataset, and the y-axis will be labeled with the range of y values. Each row in the HEATMAP dataset corresponds to one rectangle on the grid. Within that one rectangle should be printed the value of the variable RISKDIFF. If the value of RISKDIFF lies outside the margin, then an asterisk should be appended to the numeric value of RISKDIFF. A variable ASTERISK has been created in the sample code for this purpose.

The rectangle itself should be filled with color, which will be generated based on the value of RISKDIFF. Shades of red color will indicate that the TEAR rate is higher in the MYL-1601D group, while green shades will indicate a higher TEAR rate in the Novolog group. The color intensity will increase as the magnitude of RISKDIFF becomes more extreme. Colors will be programmatically assigned and will be captured in hexadecimal format in the variable COLOR in the HEATMAP dataset. This hexadecimal value can be supplied to the SAS graphics procedure.

The heat map will display the number of TEAR events in the MYL-1601D group in descending order on the x-axis and number of TEAR events in the Novolog group in descending order on the y-axis. An example is provided in the figure shells.

10.3 Laboratory Category

Laboratory Category	Parameter (SI unit)	Potentially Clinically Significant Low	Potentially Clinically Significant High
Hematology	WBC (10^9 L^{-1})	<2.0	>20.0
	Neutrophil count (10^9 L^{-1})	<1.0	NA
	Hemoglobin (g/L)	<80	>200
	Platelets (10^9 L^{-1})	<50	>999
Biochemistry	ALT (U/L)	NA	>3XULN
	AST (U/L)	NA	>3XULN
	ALK phosphatase (IU/L)	NA	>3XULN
	Total Bilirubin (umol/L)	NA	>2XULN
	Glucose (mmol/L)	<3.1	>15.0
	Creatinine (umol/L)	NA	>1.5XULN
	BUN (mmol/L)	NA	>21.4
	Sodium (mmol/L)	<130	>155
	Potassium(mmol/L)	<3.0	>6.0
	Total Protein(g/L)	NA	>95
	Albumin(g/L)	<20	NA

Laboratory Category	Parameter (SI unit)	Potentially Clinically Significant Low	Potentially Clinically Significant High
	Calcium (mmol/L)	<1.8	>3.1
Lipids	Triglycerides (mmol/L)	<0.1	>10.2
	Total Cholesterol (mmol/L)	<0.5	>10.3
	LDL Cholesterol (mmol/L)	<0.5	>9.0
	HDL Cholesterol (mmol/L)	<0.5	NA

Variable	Unit	Low	High
SBP	mm Hg	≤ 90 mmHg AND change from baseline ≤ -20 mmHg	≥ 180 mmHg AND change from baseline ≥ 20 mmHg
DBP	mm Hg	≤ 50 mmHg AND change from ≤ -15 mmHg	≥ 105 mmHg AND change from baseline ≥ 15 mmHg
Heart rate	Bpm	≤ 50 bpm AND change from baseline ≤ -15 bpm	≥ 120 bpm AND change from baseline ≥ 15 bpm
Body temperature	˚C	NA	≥ 38.3 ˚C AND change from baseline ≥ 1.1 ˚C
Weight	Kg	percentage change from baseline ≤ -7.0 %	percentage change from baseline ≥ 7.0 %