

Clinical Development

LAF237/Galvus

Clinical Study Protocol CLAF237A23156 / NCT01528254

A 5-year study to compare the durability of glycemic control of a combination regimen with vildagliptin & metformin versus standard-of-care monotherapy with metformin, initiated in treatment-naïve patients with type 2 diabetes mellitus

RAP Module 3 – Detailed Statistical Methodology

Author: [REDACTED], Novartis Trial Statistician
[REDACTED], [REDACTED] Trial Statistician

Document type: RAP Documentation

Document status: Final V3.0

Release date: 7 May 2019

Number of pages: 41

May not be used, divulged, published or otherwise disclosed
without the consent of Novartis

Document History – Changes compared to previous version of RAP module 3.

Version	Date	Changes
2.2	28NOV2018	<p>Section 2.1.5 - Background and demographic characteristics – baseline HbA1c categories were updated to be <7% and >=7%.</p> <p>2.2.1 - Definition of geographic region for efficacy analyses – geographic region categories were added as per Novartis decision.</p> <p>2.2.2.3 – Rate of loss in glycemc control over time – time (of HbA1c measurements, in years) added as random effect to the analysis model.</p> <p>2.3.1 - Overall experience of adverse events (AEs) – statement added that AEs will be analyzed by study period as well in addition to overall (including study events from study period 1, 2 and 3) in line with DMC analyses.</p> <p>Section 2.3.1.1 – statement added that AEs will be analyzed by study period for all subsection of AEs leading to discontinuation.</p> <p>Section 2.3.5 - Hypoglycemic events – statement added that events will be analyzed by study period as well in addition to overall in line with DMC analyses. Added new category to plasma glucose level list.</p> <p>Section 2.8 – Added statement that time will be used as both a fixed and random effect in analyses of rate of loss of glycemc control.</p> <p>Section 2.8 - Changes in the conduct of the study or planned analyses – it was added that the definition of initial treatment failure is changed to regard the date of the second of the two consecutive measurements of HbA1c values >=7%.</p>
3	06May2019	<p>Section 1.4 Changed duration derivation from weeks to Months.</p> <p>Section 1.6 Added the list of AEs of special interest</p> <p>Section 2.1.2 Added information of efficacy parameters to be considered for deviation code N02 and N03.</p> <p>Section 2.1.5 Removed definition of baseline HbA1c and FPG as it was redundant. Baseline is already defined in section 1.1</p> <p>Section 2.1.5 redefined age classification based on tertiles</p> <p>Section 2.1.6 redefined exposure categories in months instead of weeks.</p> <p>Section 2.2.1 Added table mentioning grouping for geographical regions</p> <p>Section 2.2.2 Removed loss of glycemc control as primary endpoint; provided additional detail to used confirmation visit for time to initial treatment failure; Since we have only one primary endpoint, removed multiple testing methods section</p> <p>Section 2.2.2.4 Added time to first treatment failure as supportive analysis; modified age categories for subgroup analysis; added smoking status, region and beta cell functional and insulin resistance as subgroups of interest; removed ethnicity</p> <p>Section 2.2.3 Added rate of glycemc control over time as secondary endpoint; [REDACTED]</p>


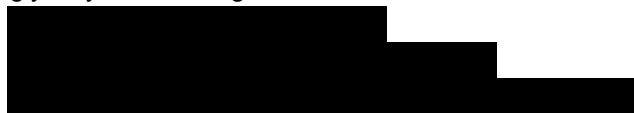
Version	Date	
		<p>analyzed and it is outside the scope of the SAP</p> <p>Section 2.2.3.3 Added the analysis of rate of glycemic control over time in this section</p> <p>Section 2.3 Removed analysis on AEs of run-in failure patients.</p> <p>Section 2.3.5 Added clarification on using chi square test to compare events only if we have sufficient observed events.</p> <p>Section 2.7 Updated power calculations based on single primary endpoint and updated yearly dropout rate.</p>

Table of contents

Clinical Development	1
List of abbreviations	6
1 Definitions and general strategy	7
1.1 Baseline definitions	7
1.2 Last On-Treatment Observation (Endpoint Observation)	7
1.3 General data handling strategy	8
1.3.1 Output data precision	8
1.3.2 Efficacy Assessment Window / Week X Measurement	8
1.3.3 Laboratory visit windows.....	9
1.3.4 Multiple laboratory measurements within same day	10
1.4 Duration of exposure to study medication.....	10
1.5 Treatment-Emergent Adverse Events.....	10
1.6 Other clinically significant events (predefined AE risks).....	10
1.7 Hypoglycemic Event Classifications.....	11
1.8 Laboratory abnormalities.....	12
1.8.1 Clinically Significant Laboratory Abnormalities.....	12
1.8.2 Hepatic enzyme and CPK elevations	12
1.9 Notable Vital Sign Abnormalities	13
1.10 GFR formula and definition.....	14
2 Statistical and analytical plans.....	15
2.1 Subjects and treatments	15
2.1.1 Analysis sets.....	15
2.1.2 Protocol Deviations and other criteria leading to exclusion of patients in analysis sets	15
2.1.3 Patient Disposition	16
2.1.4 Groupings for analysis	17
2.1.5 Background and demographic characteristics.....	17
2.1.6 Study medication.....	18
2.1.7 Concomitant medication	19
2.2 Efficacy evaluation	19
2.2.1 Definition of geographic region for efficacy analyses.....	19
2.2.2 Primary efficacy variables.....	20
2.2.3 Secondary efficacy variables.....	22
2.3 Safety evaluation	29
2.3.1 Overall experience of adverse events (AEs)	29

2.3.2	Laboratory values.....	31
2.3.3	Vital signs.....	32
2.3.4	Electrocardiograms (ECGs)	32
2.3.5	Hypoglycemic events	33
	[REDACTED]	34
	[REDACTED]	34
	[REDACTED]	34
	[REDACTED]	35
	[REDACTED]	35
2.5	Interim analyses	36
2.6	Other topics.....	37
2.7	Determination of sample size	38
2.8	Changes in the conduct of the study or planned analyses	39
References	41

List of abbreviations

AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATC	anatomical therapeutic chemical classification system
AUC	area under the curve
bid	bis in die (twice a day)
BMI	body mass index
BUN	Blood urea nitrogen
CPK	phosphocreatine kinase
CSR	clinical study report
ECG	electrocardiogram
eCRF	electronic case report/record form
EOS	end of study
FAS	full analysis set
FPG	fasting plasma glucose
GFR	glomerular filtration rate
HbA1c	glycosylated hemoglobin
	
ISR	insulin secretion rate
ISR/G	insulin secretion rate/glucose
LOCF	last observation carried forward
MDRD	modification of diet in renal disease
MedDRA	medical dictionary for regulatory activities
OGIS	oral glucose insulin sensitivity
PD	protocol deviation
PK	pharmacokinetics
PPS	per protocol set
RAN	randomized set
SAE	serious adverse event
SAF	safety set
SCR	screened-only set
SI	standard international
SMQ	standardized MedDRA queries
SYE	subject year exposure
T2DM	type-2 diabetes mellitus
ULN	upper limit of normal
VAS	visual analogue scale
WBC	white blood cell
WHO	world health organization

1 Definitions and general strategy

1.1 Baseline definitions

Unless specified otherwise, the baseline will be the measurement obtained on Day 1, or the closest sample obtained at an earlier visit (scheduled or unscheduled) to Day 1, if the Day 1 measurement is missing.

For derived variables [REDACTED] the baseline will be derived using measurements obtained on Day 1, or the closest samples obtained at an earlier visit (scheduled or unscheduled) to Day 1, if the Day 1 measurements are missing. Details of imputation methods to use if one or more of the individual measurements used in the derivation is missing, are detailed in [section 2.2.3.1](#).

Day 1 will be defined as the first day of study treatment (the earliest vildagliptin/placebo date as recorded on the vildagliptin/placebo dosage administration page of the eCRF).

Vital Signs

For blood pressure and pulse, baseline will be defined as the average of all pre-treatment visits, regardless of being scheduled or unscheduled. [REDACTED]

Routine ECG Measurements

ECGs are analyzed and interpreted locally for this study, hence only qualitative ECG measurements are recorded. ECGs are scheduled to be measured at screening (and not Day 1). Baseline is therefore defined as the measurement obtained at screening, unless there is a later, unscheduled measurement on or before Day 1 (in which case the latest of these will be used).

1.2 Last On-Treatment Observation (Endpoint Observation)

For specified efficacy variables, clinical laboratory variables, vital signs and ECG findings, the measurement obtained at the last post-baseline study visit will be considered the last on-treatment (endpoint) observation, regardless of whether it is obtained at a scheduled or unscheduled visit, up to the last scheduled visit.

For specified efficacy variables, clinical laboratory variables and ECG findings, end of Period 1 and end of Period 2 endpoints will also be of interest. The end of Period 1 endpoint is defined as the final available post-randomization assessment obtained at any visit (scheduled or unscheduled), prior to or at initial treatment failure, up to the last scheduled study visit. The end of Period 2 endpoint is defined as the final available post-randomization assessment

obtained at any visit (scheduled or unscheduled), prior to or at the initiation of insulin therapy, up to the last scheduled study visit.

1.3 General data handling strategy

1.3.1 Output data precision

Decimal places to be used in displays of demographic, background and duration of exposure variables will be as follows:

- 3 decimal places (p-values; if p-value is less than 0.001, display <0.001),
- 2 decimal places (standard errors and standard deviations),
- 1 decimal place (means, medians),
- 1 decimal place (minimums, maximums),
- 1 decimal place (percentages),
- if percentage = 100, no decimal is required.

Decimal places for efficacy and other safety summary tables and listings will be as follows:

- 3 decimal places (p-values; if p-value is less than 0.001, display <0.001),
- data precision + 2 decimal places (standard errors and standard deviations),
- data precision + 1 decimal place (means, medians),
- same as data precision (minimums, maximums),
- 1 decimal place (percentages),
- if percentage = 100, no decimal is required.

Decimal places for the Hematology, Biochemistry and Urinalysis tables will be as follows:

- 2 decimal places for standard errors and standard deviations (Hematology and Biochemistry), 3 decimal places for standard errors and standard deviations (Urinalysis),
- 2 decimal places for means, medians (Hematology and Biochemistry), 3 decimal places for means, medians (Urinalysis),
- 2 decimal places for minimums, maximums (Hematology and Biochemistry), 3 decimal places for minimums, maximums (Urinalysis).

For vital signs tables of change from baseline, 2 decimal places for mean, median, standard deviation, minimum and maximum will be used.

In outputs containing estimates and p-values for inferential purposes (e.g. ANCOVA), means, standard errors, standard deviations and confidence intervals will be output to 2 decimal places for all parameters. P-values will be output to 3 decimal places; if a p-value is less than 0.001, it will be displayed as <0.001.

1.3.2 Efficacy Assessment Window / Week X Measurement

For all efficacy variables, the measurements obtained at the scheduled visit corresponding to *X* weeks (e.g. 13, 26, ..., 260 weeks) of treatment as defined in the assessment schedule table [Tables 6-1 and 6-2 in the protocol] will be considered as Week *X* measurement (e.g. Week 13, 26, ... , 260 measurements) if $-28 \text{ days} < (\text{visit date} - \text{Week X date}) < 28 \text{ days}$. The Week *X* date is the theoretical date when the visit should have occurred (e.g. if the scheduled visit is

visit 5, corresponding to week 26, then the Week 26 date is the treatment start date + 26 weeks).

Discontinuation visits will be assigned to the closest unallocated visit. For example, if a patient discontinues at visit 4, this discontinuation visit will be re-assigned from visit 777 to visit 4 (and not visit 23). This re-assigned visit value will be stored in a separate variable, i.e. the original value will not be overwritten.

Efficacy data obtained at visits outside the above defined window or at unscheduled visits will be excluded from the analyses and tabulation of summary statistics by visit. However, they will be displayed in the patient data listings. If the last on-treatment observations of the efficacy variables are obtained at a visit outside the above defined windows or at an unscheduled visit, they will still be used for the FAS analysis based on the last observation carried forward (LOCF) imputed endpoint (see [section 1.2](#)).

Note that prior to implementation of these windows, the remapping of the efficacy laboratory parameters will be implemented.

1.3.3 Laboratory visit windows

Safety lab data

The following will be implemented for non-efficacy laboratory data:

- Safety lab data will be summarized and listed as reported on LRS panel, no additional visit windows will be used.
- Laboratory "Visit dates" from the VIS panel on eCRF will not be listed, instead laboratory listings will present the date and day of the sample collection (obtained directly from LRS).

Efficacy laboratory data

The idea of the visit window is to ensure that as many scheduled visits as possible are included in summaries of efficacy parameters over time.

For all *efficacy laboratory summaries over time* (tables and figures), a "re-defined scheduled visit" will be created as follows:

- If scheduled visit laboratory sample already exists on database then that sample remains as the "re-defined scheduled visit".
- If scheduled visit laboratory sample is missing, then search for unscheduled sample closest to actual scheduled visit date (theoretical date visit should have occurred based on treatment start date). If unscheduled sample closest to actual scheduled visit date is within 7 days of actual scheduled date then map this unscheduled visit to the scheduled visit number in the "re-defined scheduled visit". If there is a tie between unscheduled samples before and after the actual visit, take the later measurement. Note that the randomization visit laboratory measurement will not consider measurements from unscheduled visits taken after Day 1.

All efficacy outputs will include the following footnote: *"In the case of a missing scheduled visit sample, the closest unscheduled visit within 7 days of scheduled visit is used."*

1.3.4 Multiple laboratory measurements within same day

It is possible that 2 sets of laboratory measurements may be taken with the sample date. All such measurements will be included in the summaries and listings of notably abnormal or percent change from baseline summaries. For the analysis of “persistent” events (where consecutive measurements are checked) or for the analysis of parameter over time in the following key safety parameters, only the highest within-day value will be included:

- ALT
- AST
- Bilirubin (direct and/or total)
- CPK

Specifically, in the case of duplicate measurements with same sample date, only the maximum value will be selected. All samples will be listed along with sample date.

1.4 Duration of exposure to study medication

Treatment start date is defined as the earliest vildagliptin/placebo date as recorded on the vildagliptin/placebo dosage administration page of the eCRF).

Overall duration (in months) on study medication will be computed as follows:

If a complete last study drug date is available:

$$(\text{last study drug date} - \text{treatment start date} + 1) / 30.4375,$$

or if the last study drug date is missing or incomplete:

$$(\text{last visit date} - \text{treatment start date}) / 30.4375,$$

or if a patient received no dose of randomized double-blind medication:

overall duration is set to 0.

The study drug end date is defined as the last treatment date. The last treatment date in the study completion panel will be used for this study drug end date.

1.5 Treatment-Emergent Adverse Events

A treatment-emerging adverse event is defined as

- an undesirable sign, symptom, or medical condition with onset after the first dose of study medication or,
- an increase in severity of an event that is present during the pre-treatment period,
- all events included on the database will be included in AE summary tables/listings.

1.6 Other clinically significant events (predefined AE risks)

For this project, there are some predefined adverse event risks that are considered of special interest defined as below:

Important identified risks

- Transaminase elevations and Drug-induced liver injury (DILI)
- Angioedema

- Acute pancreatitis
- Skin lesions (pemphigoids)
- Hypoglycemia
- Lactic acidosis (metformin)

Important potential risks

- Serious infections
- Cardiac events in CHF (NYHA Functional Class III) patients (adjudication)
- Muscle events/ myopathy/rhabdomyolysis
- Neuropsychiatric events
- Breast cancer
- Pancreatic cancer

1.7 Hypoglycemic Event Classifications

Patients were educated to record hypoglycemic symptoms and treatment in a study diary. Study diary data were then entered in the Glycemia Study Diary or Adverse Event eCRF, respectively according to the following criteria. The Glycemia study diaries were reviewed by the study staff at each visit and assessed according to [Table 1-1](#). Note that if glucose is recorded as whole blood on the eCRF, rather than plasma, then it will be converted to a plasma value in the derived dataset, using a conversion factor of 1.12 (i.e. plasma glucose = blood glucose x 1.12).

Table 1-1 Recording and classification of glycemia study diary data

Severity	Symptoms	Plasma glucose	Action taken	Classification
Patient able to initiate self-treatment if necessary	Symptoms suggestive of hypoglycemia	< 3.1mmol/L (56mg/dL)	Enter in Glycemia Study Diary eCRF	Hypoglycemic event, grade 1
	Symptoms suggestive of hypoglycemia	≥ 3.1mmol/L(56mg /dL)	Enter in Glycemia Study Diary eCRF	Adverse Event
	Symptoms suggestive of hypoglycemia	Not taken	Enter in Glycemia Study Diary eCRF	Adverse Event
	Asymptomatic	< 3.1mmol/L(56mg /dL)	Enter in Glycemia Study Diary eCRF	Asymptomatic low blood glucose
Patient is unable to initiate self-treatment and requires assistance of another person or hospitalization	Symptoms suggestive of hypoglycemia	< 3.1mmol/L(56mg /dL)	Enter in Glycemia Study Diary eCRF & report as SAE	Hypoglycemic event, grade 2 & SAE
	Symptoms suggestive of hypoglycemia	Not taken	Enter in Glycemia Study Diary eCRF & report as SAE	Suspected hypoglycemic event, grade 2 & SAE

1.8 Laboratory abnormalities

Laboratory data will be presented in SI units only.

1.8.1 Clinically Significant Laboratory Abnormalities

Notable ranges for selected laboratory parameters of interest as listed in [Table 1-2](#) are the basis for the central laboratory to generate notable reports to alert the investigators. The notable ranges will form the basis to evaluate the incidence of clinically significant laboratory abnormalities.

Table 1-2 Notable ranges (absolute) criteria of laboratory variables of specific interest

Lab test description	SI Unit	Notable ranges	
		Low	High
Hematology			
Total leucocytes (WBC)	10E9/L	≤ 2.8	
Total leucocytes (WBC)	10E9/L		≥ 16.0
Platelet count (direct)	10E9/L	≤ 75	
Platelet count (direct)	10E9/L		≥ 700
Hemoglobin	g/L	≤115 (m) ≤ 95 (f)	–
Hematocrit	1	<0.37 (m) ≤ 0.32 (f)	–
Eosinophils	%	–	≥ 14
Biochemistry			
Sodium	mmol/L	≤ 125	
Sodium	mmol/L		≥ 160
Potassium	mmol/L	≤ 3	
Potassium	mmol/L		≥ 6
Blood urea nitrogen (BUN)	mmol/L	–	≥ 14.28
Creatinine	μmol/L	–	≥ 176.8

1.8.2 Hepatic enzyme and CPK elevations

In addition to the notable ranges, the incidence of treatment emergent hepatic enzyme elevations is of interest. These are described in [Table 1-3](#).

Treatment emergent is defined as not meeting the criterion at any pre-baseline visit, but then meeting the criterion at a post-baseline visit. Consequently a patient may be classified treatment emergent ≥5ULN even though they were not classified as treatment emergent ≥3ULN (if had a result ≥3ULN at a pre-baseline visit).

Table 1-3 Treatment emergent hepatic enzyme and CPK elevations of interest

ALT or AST ≥ 3*ULN

ALT or AST \geq 5*ULN
ALT or AST \geq 8*ULN
ALT or AST \geq 10*ULN
ALT or AST \geq 20*ULN
ALT or AST \geq 3*ULN and Bilirubin[^] \geq 1.5*ULN
ALT or AST \geq 3*ULN and Bilirubin[^] \geq 2*ULN
Bilirubin[^] \geq 2*ULN
Bilirubin[^] \geq 3*ULN
CPK \geq 5*ULN
CPK \geq 10*ULN
ALT \geq 3*ULN
ALT \geq 5*ULN
ALT \geq 8 x ULN
ALT \geq 10*ULN
ALT \geq 20*ULN
ALT \geq 3*ULN and Bilirubin[^] \geq 1.5*ULN
ALT \geq 3*ULN and Bilirubin[^] \geq 2*ULN
AST \geq 3*ULN
AST \geq 5*ULN
AST \geq 8 x ULN
AST \geq 10*ULN
AST \geq 20*ULN
Alkaline Phosphatase \geq 1.5*ULN

[^] If only direct bilirubin is measured then direct bilirubin will be used for all criteria involving bilirubin. If both total and direct bilirubin are measured (i.e. are in the laboratory specifications) throughout a study, total bilirubin will be preferred. The clinical study team has agreed that total bilirubin should be used for this study. Note that if at any visit total bilirubin has not been measured then the elevation will not be assessed for that visit (even if direct bilirubin has been measured).

In addition to the overall summary of elevations defined above, the same criteria will also be presented, restricted to “persistent elevations” only. Persistent elevations are those which meet the criterion at consecutive on-treatment measurements or at last on-treatment visit (this includes assessments within one day of the last dose of study treatment). The purpose of the persistent definition is to remove transient elevations from the summary (i.e. those elevations which appeared at one visit but normalized at the next visit).

1.9 Notable Vital Sign Abnormalities

Notable vital sign abnormalities [REDACTED] are defined in [Table 1-4](#).

Table 1-4 Notable vital sign abnormalities

Vital signs	Notable abnormalities		
	Low	High	
Pulse (beats/min)	either ≤ 50 + decrease $\geq 30^*$ or < 40	either ≥ 120 + increase $\geq 25^*$ or > 130	
Blood pressure (mmHg)	Systolic	either ≤ 90 + decrease $\geq 30^*$ or < 75	either ≥ 180 + increase $\geq 30^*$ or > 200
	Diastolic	either ≤ 50 + decrease $\geq 20^*$ or < 40	either ≥ 105 + increase $\geq 20^*$ or > 115
██████████	████████████████████	████████████████████	

* Refers to post-baseline value as compared to baseline value

The notable abnormality of ‘ ≥ 120 + increase ≥ 25 or > 130 ’ will be interpreted as that the value in the treatment period is ≥ 120 and the increase from baseline is ≥ 25 , or the value in the treatment period is > 130 .

1.10 GFR formula and definition

The following formula is used to estimate Glomerular Filtration Rate (GFR) modification of diet in renal disease (MDRD) (Rigalleau 2005):

$$\begin{aligned}
 \text{GFR}_{\text{MDRD}} (\text{mL/min}) \times (1.73/\text{m}^2) = & \\
 & 175 \times (\text{serum creatinine } \{\text{mg/dL}\})^{-1.154} \\
 & \times (\text{age } \{\text{years}\})^{-0.203} \\
 & \times 0.742 \text{ (if female)} \\
 & \times 1.210 \text{ (if race is Black)}
 \end{aligned}$$

GFR will be available in the raw laboratory data (and used to assess inclusion/exclusion criteria) but will be re-derived within the derived datasets using this RAP-specified formula. This will ensure consistency with other studies within the program.

Baseline GFR (MDRD) is a derived variable calculated using the baseline serum creatinine, ██████████ & age value, i.e. the value obtained Day 1, or the sample obtained at an earlier visit (scheduled or unscheduled) which is closest to Day 1, if the Day 1 measurement is missing. Hepatic function is categorized as follows:

- Normal: GFR (MDRD) > 80
- Mild: GFR (MDRD) ≥ 50 - GFR (MDRD) ≤ 80
- Moderate: GFR (MDRD) ≥ 30 - GFR (MDRD) < 50
- Severe: GFR (MDRD) < 30

2 Statistical and analytical plans

Data will be analyzed [REDACTED] according to the data analysis [section 9 of the study protocol](#) which is available in [Appendix 16.1.1 of the CSR](#). Important information is given in the following sections and details are provided, as applicable, in [Appendix 16.1.9 of the CSR](#).

2.1 Subjects and treatments

2.1.1 Analysis sets

Screened-only set (SCR): The SCR consists of all patients who were screen failed after the first visit or who entered the run-in phase but were not randomized. Except for the listing of individual patients in this SCR with the reasons for not being randomized, and the tabulation of patients in this SCR by discontinuation reason, no other analysis will be performed on this analysis set.

Randomized set (RAN): The RAN consists of all randomized patients.

Full analysis set (FAS): The FAS consists of all randomized patients who receive at least one dose of randomized study medication (vildagliptin or placebo) and have at least one post-randomization assessment of any efficacy parameter. Following the intent-to-treat principle, patients will be analyzed according to the treatment approach they are assigned to at randomization.

Safety set (SAF). The SAF consists of all patients who receive at least one dose of randomized study medication (vildagliptin or placebo). Patients will be analyzed according to the treatment approach received. If a patient receives both vildagliptin and placebo in Period 1, then the patient will be included in the vildagliptin group. Note that the SAF allows the inclusion of non-randomized patients who received the study drug in error.

Per Protocol set (PPS): The PPS is a subset of FAS and consists of all randomized patients who receive at least one dose of randomized study medication (vildagliptin or placebo), have at least one post-randomization assessment of any efficacy parameter in Period 1, do not discontinue the study prior to Week 26, and have no major protocol deviations occurring during Period 1.

2.1.2 Protocol Deviations and other criteria leading to exclusion of patients in analysis sets

Patients will be excluded from analysis based on the protocol deviations (PDs) in [VAP Module 3](#) and other criteria described in [Table 2-1](#). The PDs described in [VAP Module 3](#) will be identified manually by field monitors/CTH and/or via PD programming, and will be provided in the VIOPTO source dataset, whilst those in ([Table 2-1](#)) will be programmed by Biostatistics using source data, and stored in the derived dataset NOVIOPTO. The severity codes assigned to each deviation as described in [VAP Module 3](#) and [Table 2-1](#) will be used to classify patients as described in [Table 2-2](#). Note that other protocol deviations will also be identified. However, patients will not be excluded from analysis based on these other deviations (with severity code 49) including violations of entry criteria on prior use of anti-

diabetic agents and BMI range. The subject classification as specified in [Table 2-2](#), will be defined based on severity codes prior to un-blinding the database with respect to treatment.

Table 2-1 Other criteria leading to exclusion of patients in analysis sets

Description of Deviation	Deviation code	Severity code	Inclusion/Exclusion in Analyses
No dose of randomized, double-blind medication (vildagliptin or placebo) taken during Period 1	N01	8	Excluded from all efficacy and safety analyses
No post-randomization efficacy assessment during the treatment period (Periods 1, 2 or 3)	N02	0	Excluded from efficacy analyses
No post-randomization efficacy assessment during Period 1	N03	1	Excluded from PPS analyses
Patient discontinued from the study prior to Week 26	N04	1	Excluded from PPS analyses

Subjects with no post-randomization efficacy assessment of Hba1c during the treatment period (Periods 1, 2 or 3) will be excluded from efficacy analyses (Deviation code N02). Subjects with no post-randomization efficacy assessment of Hba1c, fasting insulin, fasting glucose, fasting C-peptide and meal-test parameters during Period 1 will be excluded from PPS analyses (Deviation code N03).

Table 2-2 Subject classification

Analysis set	Deviation Severity Codes			
	0	1	8	49
Randomized Set (RAN)	Yes	Yes	Yes	Yes
Full Analysis Set (FAS)	No	Yes	No	Yes
Safety Set (SAF)	Yes	Yes	No	Yes
Per Protocol Set (PPS)	No	No	No	Yes

Note: yes/no means whether patients with PDs or non-PD criteria of a specific code are included (yes) or excluded (no) from an analysis set.

The number and percentage of patients in the RAN with protocol deviations and other criteria leading to exclusion of patients from analysis sets will be tabulated by treatment approach. Protocol deviations that do not lead to exclusion from analysis populations will be further summarized for the RAN by treatment approach. Summaries will be repeated by Period.

2.1.3 Patient Disposition

The number and percentage of randomized patients (RAN) who completed the study and who discontinued will be tabulated by treatment approach. Discontinuation will be further broken down by reason in the tabulation. Summaries will be repeated by Period, including information about the number of patients continuing into the next Period.

Completion information will be listed for the RAN.

Patients who were screened for the trial and entered the run-in phase but were not randomized (SCR) will be tabulated and listed with reasons for not being randomized. These summaries will include both screen failures and run-in failures (i.e. patients who discontinue within the screening period and patients who discontinue within the run-in period). Note that the patients will be presented together and not separated out into separate groups for screen failures and run-in failures.

2.1.4 Groupings for analysis

The number and percentage of patients in each analysis set will be summarized by treatment approach. The number and percentage of patients included in the meal test, [REDACTED] subsets will also be summarized by treatment approach. Meal test [REDACTED] count data will be presented as part of the efficacy evaluations. [REDACTED]

2.1.5 Background and demographic characteristics

Demographic, background data and key efficacy variables at baseline will be summarized for all randomized patients (RAN) by the treatment approach that patients are assigned to at randomization. Categorical variables gender, race, ethnicity, baseline age group, current smoking status, baseline GFR (MDRD) category, baseline body mass index (BMI) group, and baseline HbA1c group will be summarized by frequency and percentage. The continuous variables baseline age, height, [REDACTED], baseline BMI, duration of T2DM, baseline HbA1c and baseline FPG will be summarized by mean, standard deviation, median, minimum and maximum.

All baseline and background characteristics (except HbA1c, FPG and GFR) will be summarized at study entry (screening visit) or at the randomization visit if screening (Week - 5) measurements are not collected.

Baseline body mass index (BMI) will be calculated (i.e. not taken from the eCRF) from the height and weight measured at screening (Week -5), $BMI = \text{weight} / \text{height}^2$, where weight is in kilograms (kg) and height is in meters (m).

The classification of the baseline HbA1c, BMI, age category and GFR (MDRD) category are based on the following definitions:

- HbA1c category (< 7% and $\geq 7\%$)
- BMI (< 30 kg/m² and ≥ 30 kg/m² at Visit 1)
- Age at Visit 1 by tertiles i.e. <48 years, 48 - < 62 years and ≥ 62 years)
- GFR (MDRD) category (Normal: >80 mL/min, Mild: ≥ 50 mL/min - ≤ 80 mL/min, Moderate: ≥ 30 mL/min - <50 mL/min, Severe: 0 mL/min - <30 mL/min)

Baseline comparability across treatment approaches for all randomized patients (RAN) will be examined in a [Section 16.1 listing](#) using a chi-square test for the categorical variables, and a two-sample t-test for the continuous variables. The p-values are provided for descriptive purposes, and will not to be considered as the formal basis for determining factors to be included in the statistical analysis model.

In addition, all relevant medical history and history of diabetes and complications will be summarized by primary system organ class, preferred term and treatment approach using frequency tables in Post-text tables for all RAN patients.

2.1.6 Study medication

The study consists of three treatment periods.

In Period 1 patients will be randomized to metformin + vildagliptin 50mg bid, or metformin + placebo. If initial treatment fails to maintain HbA1c < 7.0%, confirmed at two consecutive scheduled study visits (after remapping in [section 1.3.3](#) has been applied), starting from Visit 4 (Week 13), then vildagliptin will be added to metformin in the metformin-only arm in Period 2. The Period 1 end date is defined as the date of the second of the two consecutive scheduled study visits, i.e. the initial treatment failure date. The Period 2 start date is defined as the next day.

In Period 3, insulin is considered based on local guidelines. Both patients and investigators will remain masked to the treatment allocation in Period 1, and the study will compare the two different treatment approaches. The Period 3 start date is defined as the date of insulin initiation. The Period 2 end date is defined as the previous day.

Treatments are referred to as ‘Treatment approach’ throughout this analysis plan to acknowledge that patients are not on the same treatment for the duration of the study. When summaries are detailed for the ‘Treatment Period’, they will include data from Periods 1, 2 and 3, unless specified otherwise.

The duration of exposure to study medication (in months) during the treatment period will be computed for all patients using the algorithm outlined in [Section 1.4](#) and summarized by treatment approach for the RAN both descriptively (i.e., mean, standard deviation, median, minimum and maximum) and by 3 - monthly intervals (0 - < 3 months, 3 - < 6 months, ..., ≥ 60 months).

The study Period 1 duration (in months) will be also similarly summarized for the RAN. The study Period 1 duration is defined as (initial treatment failure date – treatment start date + 1)/30.4375 for patients who entered study Period 2. The study Period 1 duration is the same as overall study duration for patients who did not enter into Period 2.

The study period from treatment start to the end of Period 2 will be computed for all patients who initiate insulin in Period 3 and for patients discontinuing the study in Period 2 due to being unable or unwilling to initiate insulin for treatment intensification in Period 3. The duration will be defined as (initiation date of insulin – treatment start date)/30.4375 for patients who initiate insulin, and (last study drug date – treatment start date + 1)/30.4375 for patients who do not enter Period 3.

Total subject year exposure (SYE) will also be summarized by treatment approach for the RAN. SYE (years) for each treatment will be calculated as (duration of exposure in days, disregarding any treatment interruptions, for all subjects in that treatment approach) / 365.25. For patients who did not take any study medication, the exposure is set to zero. This summary will be repeated on Period 1 data only, and also on Vildagliptin exposure only.

Study medication information will be listed for the RAN, including information about vildagliptin, placebo, metformin and insulin use.

2.1.7 Concomitant medication

Concomitant therapies taken during the treatment period will be summarized by anatomical therapeutic chemical classification system (ATC) class and preferred term according to the treatment approach patients are assigned to at randomization in the RAN. The number and percentage of randomized patients (RAN) in each class and preferred term will be tabulated by treatment approach. This summary will be repeated on period 1 data only, i.e. concomitant therapies taken during Period 1.

Use of anti-diabetic medications prior to study entry (as recorded on the eCRF) will be summarized by ATC class, preferred term and treatment approach.

The duration and average daily dose of metformin taken during the screening and run-in period will be summarized descriptively (mean, standard deviation, median, minimum and maximum), using data from 35 days prior to screening through to 35 days after randomization. This will be based on information collected on the i) Metformin dosage administration page of the eCRF (considering only dosages which are at most 35 days after randomization); and ii) information collected on the prior anti-diabetic medication eCRF page with a preferred term containing 'METFORMIN' and a dose of 'mg' (considering only dosages which are at most 35 days prior to the screening visit). The average daily dose of metformin during this window will also be summarized categorically (<2000 mg, ≥2000 mg).

Listings of concomitant medications and prior anti-diabetic medications will be presented by treatment approach and patient number.

2.2 Efficacy evaluation

2.2.1 Definition of geographic region for efficacy analyses

Centers/countries with a small number of patients will be pooled to create geographic regions. Larger centers/countries will be considered as their own geographic region. There are six geographic regions planned: Europe, Latin America, South-East Asia, East Asia, Australia and Africa. The pooling of centers and/or countries to create geographic regions will be confirmed once all patients have been randomized. The allocation of countries to the specific geographic regions will be as follows.

Centers/Countries	Geographic region
AUT, BGR, CZE, DEU, ESP, EST, FIN, HUN, ISR, ITA, LTU, LVA, NOR, POL, ROU, RUS, SVK, TUR	Europe
ARG, BRA, COL, DOM, GTM, MEX, PAN, PER	Latin America
IND, MYS, PHL	South-East Asia
HKG, KOR, TWN	East Asia

AUS	Australia
ZAF	Africa

2.2.2 Primary efficacy variables

2.2.2.1 Variable definition

The primary efficacy variable is time to confirmed initial treatment failure. The confirmed initial treatment failure is defined as a patient with HbA1c measurements $\geq 7.0\%$ at two consecutive scheduled visits (after remapping in [section 1.3.3](#) has been applied), starting from Visit 4 (Week 13). The time to confirmed initial treatment failure is defined as the time from randomization until the time when the second of the two HbA1c measurement $\geq 7.0\%$ was determined after at least 13 weeks of treatment (Visit 4), i.e. the end of Period 1. The rationale of starting the treatment failures count from Visit 4 is that treatment guidelines consider as treatment failure patients who do not achieve HbA1c $< 7.0\%$ after 3 months of treatment ([ADA-EASD, 2009](#)).

The main study conclusion will be based on the analysis of the primary efficacy variable using the FAS. Analyses based on the PPS will also be performed to assess the robustness of the conclusion.

2.2.2.2 Handling of missing values/censoring/discontinuations

In the analysis for the primary efficacy variable of time to treatment failure, patients who discontinue the study for any reason during Period 1 (lack of efficacy, lost to follow-up, AE or abnormal laboratory values etc.) will be treated as censored values at the time of discontinuation. Patients who remain under the threshold (or whose measurement above the threshold is not confirmed at next scheduled visit) will be censored at the time of last study visit.

No imputation will be used for missing HbA1c measurements. As detailed in [section 1.3.3](#), re-defined scheduled visits will be created and so it should be the case that there is always a measurement for a scheduled visit, as an unscheduled visit should be carried out to ensure there was an HbA1c value available. However, if there is a case where the HbA1c value is missing then when looking for consecutive scheduled visits, only visits with non-missing HbA1c measurements will be considered (e.g. if Visit 4 HbA1c $\geq 7.0\%$, Visit 5 HbA1c is missing, and Visit 6 HbA1c $\geq 7.0\%$, then this will be considered treatment failure).

2.2.2.3 Statistical model, hypothesis, and method of analysis

Time to initial treatment failure

The primary statistical hypothesis of time to initial treatment failure will be assessed by a 1-sided test of superiority of the combination treatment with vildagliptin + metformin versus metformin monotherapy with α -level of 0.025 with a null hypothesis that the hazard-ratio between the combination treatment vildagliptin + metformin and metformin monotherapy is equal or greater than 1, versus the 1-sided alternative hypothesis that the hazard-ratio is less than 1:

$H_0: \lambda_{\text{vildagliptin+metformin}}/\lambda_{\text{metformin}} \geq 1$

versus

$H_a: \lambda_{\text{vildagliptin+metformin}}/\lambda_{\text{metformin}} < 1$

where $\lambda_{\text{vildagliptin+metformin}}$ and $\lambda_{\text{metformin}}$ are the hazard-rates of the initial treatment failure for vildagliptin 50mg bid + metformin up to 1000mg bid and metformin up to 1000mg bid monotherapy respectively.

The time to confirmed failure will be derived as the time from randomization to the second of two consecutive scheduled visits (after remapping in [section 1.3.3](#) has been applied), at which HbA1c $\geq 7.0\%$ is measured, starting from Visit 4 (13 weeks after randomization).

The primary efficacy analysis will use a Cox proportional hazard regression model to assess the probability of initial treatment failure, with treatment approach and geographic region as classification variables and baseline HbA1c as a covariate. The hazard-ratio and associated 95% confidence interval as well as the p-value estimated from the above model will be presented by treatment approach.

The initial treatment failure rate over time by treatment approach will be summarized using estimates and 95% confidence intervals from a Kaplan-Meier analysis. The Kaplan-Meier estimates will be plotted, with symbols to indicate censored values.

The primary analysis for this primary efficacy variable will be performed using the FAS. It will be performed in the PPS as supportive analysis as well.

The model assumptions will be checked by plotting the Schoenfeld residuals for the treatment effect obtained from the model against time. A random spread of residuals with no clear pattern will indicate that the proportional hazards assumption is valid. The results will be presented in CSR Section 16 table(s).

2.2.2.4 Supportive analyses

Time to first treatment failure

The time to first treatment failure will be derived as the time from randomization to the first of two consecutive scheduled visits (after remapping in [section 1.3.3](#) has been applied), at which HbA1c $\geq 7.0\%$ is measured, starting from Visit 4 (13 weeks after randomization).

Time to first treatment failure will be analyzed in the similar way as primary analysis.

Descriptive analysis in HbA1c over time

Summaries of absolute values and change from baseline to end of Period 1 in HbA1c by treatment approach and visit will be presented in the FAS. Figures will be produced showing the mean HbA1c and standard error by visit during the study Period 1 for each treatment approach. The summary table and figure will be repeated for PPS.

In addition, the above descriptive analysis (including summary table and figure) will be performed for FAS on HbA1c data collected up to the end of Period 2, as well as data collected during the entire duration of the study.

Descriptive subgroup analysis

Summaries of absolute values and changes from baseline to end of Period 1 in HbA1c will be presented by treatment approach. The analysis will be performed on FAS. The list of subgroups of interest is listed below:

1. Baseline HbA1c category ($<7\%$, $\geq 7\%$)
2. BMI ($<30\text{kg/m}^2$, $\geq 30\text{kg/m}^2$ at Visit 1)
3. Age at Visit 1 by tertiles i.e. < 48 years, $48-< 62$ years, ≥ 62 years
4. Gender
5. Race
6. Smoking status
7. Geographical regions

2.2.3 Secondary efficacy variables

2.2.3.1 Derivation of secondary [REDACTED] variables

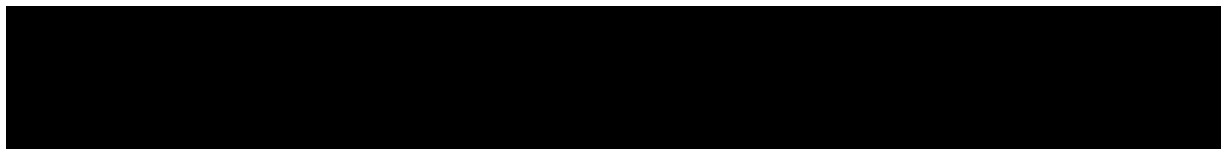
Rate of glycemc control over time

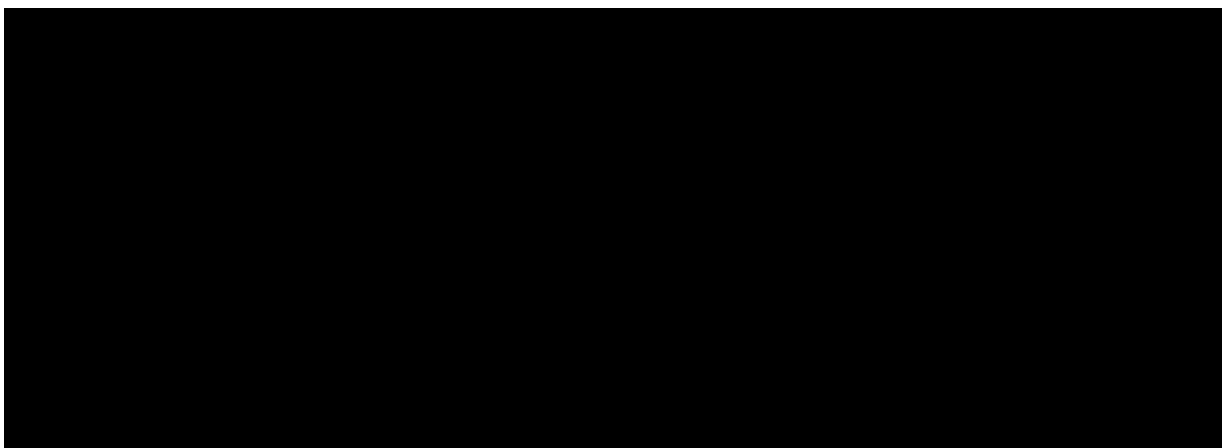
The rate of loss of glycemc control over time will be estimated by an annualized slope of HbA1c over time from Visit 5 (Week 26) to the end of Period 1. A mixed effect model including a random coefficient will be used for analysis. The rationale of choosing Visit 5 as starting point for estimating the rate (slope) of loss of glycemc control over time is based on data from study CLAF237A2307 ([Scherbaum W et al, 2008](#)) where the maximal reduction in mean HbA1c from baseline was between 24 and 32 weeks, therefore a trend for loss in glycemc control is sensible to be estimated from a similar time point.

In the analysis of the rate of loss in glycemc control over time, data used in analysis will be censored at the start of Period 2, i.e. the only data included in the model will be from Visit 5 (Week 26) to the end of Period 1.

FPG

If the FPG value (standard non meal-test parameter) is missing, the -20 min meal-test value (obtained on the same day) can be used to impute. If the -20 min value is also missing, the 0 min value will be used.





Calculation of OGIS

OGIS will be calculated as a function of glucose and insulin (meal-test laboratory parameters):

$$\text{OGIS} = 0.5 * \left\{ \begin{aligned} & [((P_5 * (G(90) - G_c)) + 1) * \text{CI_OGTT}] \\ & + \text{sqrt} \left\{ \begin{aligned} & [((P_5 * (G(90) - G_c)) + 1) * \text{CI_OGTT}] ^2 \\ & + [4 * P_5 * P_6 * (G(90) - G_c) * \text{CI_OGTT}] \end{aligned} \right\} \end{aligned} \right\}$$

where,

- $\text{CI_OGTT} = \left\{ P_4 * \left\{ (P_1 * D_0) - [V * (G(120) - G(90))] / 30 \right\} / G(90) + (P_3 / G(0)) \right\} / \left\{ I(90) - I(0) + P_2 \right\}$
- D_0 (g/m²) is the glucose dose included in the standard meal administered to patients who participate in the standard meal-test ($D_0 = (73.2 * 5.551) / \text{BSA}$ for the standard meal as planned in [Protocol Section 6.4.3](#). It will be assumed all patients receive a standard meal),
- $\text{BSA} = 0.164 * (\text{Weight in kg} ** 0.515) * ((0.01 * \text{Height in cm}) ** 0.422)$,
- $V = 10000$ (ml/m²) is the assumed glucose distribution volume,
- $G(0)$, $G(90)$ and $G(120)$ are glucose concentration values (taken from meal-test laboratory data) at timepoints 0, 90 and 120 min,
- $I(0)$ and $I(90)$ are insulin concentration values (taken from meal-test laboratory data) at timepoints 0 and 90 min,
- $P_1=6.5$, $P_2=1951$, $P_3=4514$, $P_4=792$, $P_5=0.0118$, $P_6=173$, $G_c=5$ are constants ([Mari, et al 2001](#)).

If the value of glucose or insulin values at 0 minutes are missing, then the -20 min meal-test value (obtained on the same day) can be used to impute. If the -20 min value is also missing, the fasting parameter result (i.e. standard non meal-test parameter) will be used. Both glucose

and insulin should be imputed using the same timepoint (regardless if one has value available).

No imputation will be used for missing 90 or 120 minute values.

Calculation of ISR

Separately for each patient and each meal-test, deconvolution is applied to estimating insulin secretion to assess β -cell function. C-peptide, which is co-secreted with insulin in an equimolar ratio, is preferred to insulin as a basis for estimating insulin secretion.

Mathematically, we need to solve for the insulin secretion rate r given the impulse response h , and plasma C-peptide c , where

$$c(t) = \int_{-\infty}^t h(t-s)r(s)ds.$$

The impulse response is the PK model for an intravenously administered unit bolus dose of C-peptide when the endogenous C-peptide production has been suppressed, i.e.

$$h(t-s) = (1/V)[fe^{-\alpha(t-s)} + (1-f)e^{-\beta(t-s)}],$$

where V , f , α , and β are a subject's parameters, which are generally unknown. However, from clinical studies these parameters have been estimated, and are related to a subject's disease status, body surface area, age, and gender. Thus, h becomes known.

Insulin secretion rates are modeled by a piecewise continuous linear function with knots at the actual sampling time. Changes in slopes are penalized through a single regularization parameter that is determined by a root finding method to produce similar residual standard deviation for fitted C-peptide concentrations as would be expected from assay error standard deviations given in the kit instructions.

The IMMULITE 2000 C-peptide kit instructions were used to model intra-assay SD as a function of the true C-peptide concentration $C1$ expressed in conventional units of ng/mL by simple linear regression as:

$$SD = 0.04056C1 + 0.06311\text{ng/mL}.$$

With concentration of C-peptide expressed alternatively in SI units, i.e., $C2$ expressed in nM (1ng/mL C-peptide = 0.331nM C-peptide) the equations for SD becomes:

$$SD = 0.04056C2 + 0.331 \times 0.06311\text{nM}.$$

The following ln-transformation with a shift parameter of observed and fitted C-peptide values is used to stabilize the variance in performing the deconvolution:

$$\ln(C1 + 0.06311/0.04056) \text{ or } \ln(C2 + 0.331 \times 0.06311/0.04056).$$

Nonlinear least squares with ISR constrained non-negative using PROC IML function NLPLM in SAS® is used to solve for ISR at the sampling time points using actual sampling times. The bisection method is used to iteratively determine the regularization parameter such that the sum of squares of the observed C-peptide minus the fitted C-peptide expressed in either units equals $0.040562n$ where n is the number of non-missing C-peptide concentrations in the meal-test.

Validated macros (ISR and ISR_IML2) have been provided by Novartis and will be used to derive the ISR in █████ production and validation programs.

Calculation of AUC

AUC of ISR/G is defined as ISR relative to glucose 0-2h (pmol/min/m²/mM) = AUC_{0-2hr} of ISR/AUC_{0-2hr} of glucose. (Glucose is taken from the meal-test laboratory data).

For the patients participating in the meal-test, the laboratory parameters glucose, insulin and C-peptide are also to be measured at the following 7 time-points (minutes in relation to the meal time): -20min, 0min (immediately prior to meal ingestion), 15min, 30min, 60min, 90min, 120min. Let L_1, L_2, \dots, L_7 denote the lab measurements at these 7 time-points (i.e., L_1 is the value measured at 20min pre-meal, L_2 is the value measured at 0min pre-meal, ..., L_7 is the value measured at 2h after the meal) and T_1, T_2, \dots, T_7 be the corresponding times (actual times in units of h), the AUC_{0-2hr} is calculated (using the trapezoidal rule) as:

AUC_{0-2h} =

$$[(L_2 + L_3) \times (T_3 - T_2) + (L_3 + L_4) \times (T_4 - T_3) + (L_4 + L_5) \times (T_5 - T_4) + (L_5 + L_6) \times (T_6 - T_5) + (L_6 + L_7) \times (T_7 - T_6)] / 2.$$

Note for above calculation:

- 1) L_1 and T_1 are not used in the calculation. However, if L_2 (the value at 0 min) is missing, the L_1 (the value at -20 min) will be used in its place, but T_2 should still be used in the calculation (i.e., no substitution of T_2 with T_1 is allowed). If the -20 min AND 0 min values are missing, the fasting value (i.e. the standard non meal-test value) will be used to impute the 0 min value. If the 0 minute ISR is missing due to missing 0 minute c-peptide, then the missing c-peptide value should be imputed using this approach, in order to obtain a non-missing 0 minute ISR value.
- 2) If L_i (a lab value) is not missing, but the corresponding time T_i is missing, T_i will be imputed by the planned time (or scheduled time).
- 3) If L_i is missing at a timepoint T_i , the next available value L_{i+1} and timepoint T_{i+1} will be used for calculating AUC.
- 4) AUC_{0-2h} will only be calculated for patients for whom there are at least 3 timepoints (including one of the first two timepoints (i.e., 0 or -20 min pre-meal), the last timepoint (i.e., 2 hours post-meal), and a timepoint in between) where the actual laboratory values are known.

2.2.3.2 Variable definition

The following secondary [REDACTED] efficacy variables will be analyzed. They are grouped by type of analyses for the convenience of describing the related analysis methods.

[REDACTED]
[REDACTED]

Variables related to the rate of change of β -cell function and insulin sensitivity

- Rate of loss of β -cell function (assessed using the AUC of ISR/G) from Visit 4 (week 13) to: i) end of Period 1; ii) end of Period 2; iii) EOS.
- Rate of change in insulin sensitivity (assessed using the OGIS) from Visit 4 (week 13) to: i) end of Period 1; ii) end of Period 2; iii) EOS.

Variables related to the rate of loss in glycemic control over time

- Rate of loss in glycemic control in HbA1c from 26 weeks after the start of Period 2 to end of Period 2.
- Rate of loss in glycemic control in FPG, i) from Visit 5 (week 26) to: i) end of Period 1; ii) from 26 weeks after start of Period 2 to end of Period 2.

Variables related to change from baseline to endpoint

- Change in AUC of ISR/G (as an assessment of β -cell function) from baseline to: i) end of Period 1; ii) end of Period 2; iii) EOS.
- Change in OGIS (as an assessment of insulin sensitivity) from baseline to: i) end of Period 1; ii) end of Period 2; iii) EOS.

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

2.2.3.3 Statistical analysis methods

Rate of loss in glycemic control over time

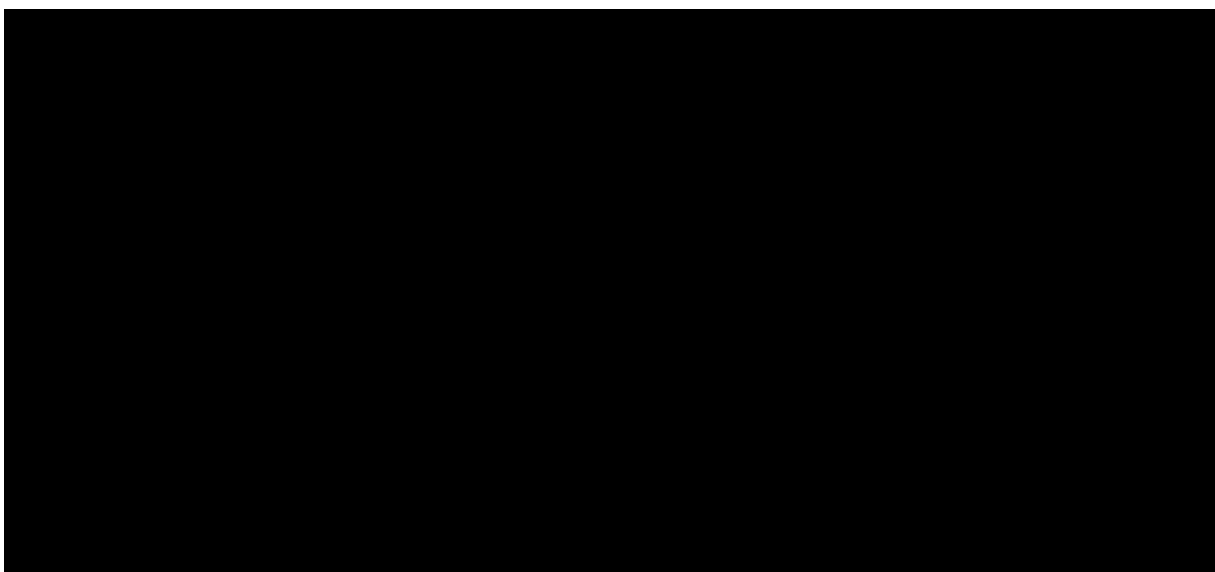
The rate of loss in glycemic control over time will be estimated by the slope of HbA1c over time (in years) as a random coefficient in a linear mixed effect model: the model will be fitted to HbA1c data collected from Week 26 and onwards, up to and including the end of the

Period 1 visit, i.e. up to and including the initial treatment failure date. It includes treatment approach, geographic region, baseline HbA1c, time (of HbA1c measurements, in years) and the interaction of treatment approach by time as the fixed effects, and time and intercept as random effects. Treatment approach and geographic region will be considered as classification variables while baseline HbA1c and time are the covariates. The unstructured covariance will be used as the covariance structure within patients. The actual hypothesis will be tested using the treatment approach by time interaction term. Note that it is not an issue that patients who 'fail' at Week 26 will only contribute one data point to the analysis, as this is expected in only a few patients.

The mean slopes within each treatment approach and the difference in mean slopes between two treatment approaches as well as the p-value obtained from the test using the above model will be presented. Graphical representation will also be produced as required.

The analysis will be performed using the FAS. It will be performed in the PPS as supportive analysis as well.

The assumptions for inference using the model as described above (i.e., errors are independently normally distributed with constant variance) will be checked using CSR Section 16 table(s) by subjective examination of the presented residual plots: normal probability plot and residuals versus fitted values plots.



Variables related to the rate of change of β -cell function and insulin sensitivity over time

The rate of change of β -cell function and insulin sensitivity (assessed using the AUC of ISR/G and OGIS respectively) from Visit 4 (week 13) to the end of Period 1, or from Visit 4 (week 13) to the end of Period 2, or from Visit 4 (week 13) to the end of study will be assessed using a similar random coefficient linear mixed effect model used for the endpoint 'rate of loss in glycemic control in HbA1c from Visit 5 (Week 26) to the end of Period 1'.

This analysis is performed in the FAS only. Data collected up to the end of Period 1, or the end of Period 2, or to the end of the study will be used in the analysis for assessing the rate of change from Visit 4 to end of Period 1, or from Visit 4 to end of Period 2, or from Visit 4 to the end of the study respectively.

A general downward trend from Visit 4 to the end of study can be assumed for these parameters.

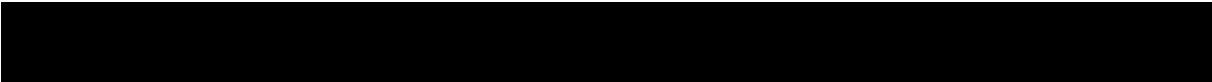
Variables related to the rate of loss in glycemc control over time

The rate of loss in glycemc control in HbA1c or FPG from 26 weeks after the start of Period 2 to the end of Period 2 will be assessed for patients who start insulin therapy in Period 3 (i.e. for treatment intensification) or discontinue the study during Period 2 due to not being able or unwilling to initiate insulin therapy in Period 3. The random coefficient linear mixed effect model used for the endpoint ‘rate of loss in glycemc control in HbA1c from Visit 5 (Week 26) to the end of Period 1’ will be used in this secondary variable. Only data collected between initiation of Period 2 to insulin initiation on patients who take insulin therapy in Period 3 (i.e. for treatment intensification) or who discontinued the study in Period 2 due to not being able or unwilling to initiate insulin therapy in Period 3 will be included in this model. Patients who complete the study in Period 1 or Period 2 will not be included in the analysis. The analysis will be performed in the FAS only. Note that it is not known how many patients will proceed into Period 3, so may need to later review whether the planned analysis for the subgroup of patients entering Period 3 (or being unable/unwilling to) is sensible.

The rate of loss in glycemc control in FPG from 26 weeks to the end of Period 1 will be assessed using the same random coefficient linear mixed effect model used for the endpoint ‘rate of loss in glycemc control in HbA1c from Visit 5 (Week 26) to the end of Period 1’. This analysis is performed in the FAS only.

Variables related to change from baseline to endpoint

An analysis of covariance (ANCOVA) model with treatment approach, geographic region as classification variables and baseline value as a covariate will be used to assess all secondary efficacy variables related to change from baseline to endpoint. The least squares mean (“adjusted mean”) change from baseline for each treatment approach, the difference in the least squares mean changes between the two treatment approaches, and the two-sided adjusted 95% confidence interval along with the p-value for the difference will be obtained from this analysis model and presented.

 The assessment may be performed within two months before or after the actual study visit for practical reasons, so the date of the ophthalmic assessment may differ from the actual study visit. An endpoint is not defined for this data.

Supportive descriptive analysis for secondary [REDACTED] variables over time

Summaries of absolute values and change from baseline to: i) end of Period 1; ii) end of Period 2; iii) EOS; in secondary [REDACTED] variables (including FPG, [REDACTED], AUC of ISR/G, OGIS, [REDACTED]) by treatment approach and visit will be presented in the FAS. [REDACTED].

2.3 Safety evaluation

The safety and tolerability of vildagliptin 50mg bid as add-on to metformin over 5-years of treatment will be compared to placebo as add-on to metformin based on data collected during the treatment period in SAF. All analyses will use all available data collected during the entire study. In addition, key safety analyses will be performed on data obtained during Period 1 only. Patients will be analyzed according to the treatment approach they received (vildagliptin 50mg bid + metformin, placebo + metformin).

The incidence of treatment emergent adverse events (events started after the first dose of randomized study medication or events present prior to start of randomized treatment but increased in severity) will be summarized by primary system organ class, preferred term, severity and relationship to study drug. Any event captured on the database will be included in AE summary tables/listings.

In order to assess safety and tolerability of vildagliptin as compared to placebo, key safety variables (overall AEs, SAEs, AEs leading to study drug discontinuation or study drug interruption, incidence of hypoglycemia, predefined AE risks, predefined categories of liver enzyme (ALT/AST) and CPK and persistent elevations) will also be summarized by treatment approach for the SAF.

2.3.1 Overall experience of adverse events (AEs)

The number and percentage of patients reporting any adverse event during the treatment period of the study will be summarized by primary system organ class, preferred term and treatment approach overall (including study periods 1, 2 and 3) and by study period. Percentages will be based on the number of patients in the SAF per treatment approach and study period. An AE will be assigned to a period if its start date is after or on the start date but not later than the end date of that period. All AEs will be assigned to a specific study period – if it is not possible to assign an AE into a specific study period of the patient based on the AE start date, then the event will be assigned to the last study period started by the patient. This will ensure that all AEs will be categorized to a specific study period

The most common adverse events reported (≥ 2 % in any group) during the treatment period of the study will be presented in descending frequency starting from the most common event in a Text Table. A complete table of all events reported in descending frequency will be prepared.

Adverse events that were suspected to be related to study drug will be summarized. Note that if the relationship to study drug is missing for any adverse event, then that adverse event will be assumed to be related to study drug.

The maximum severity of all adverse events during the treatment period of the study will be summarized. Note that if the severity is missing for any adverse event, then that adverse event will be assumed to be severe.

An overall summary of the number and percentage of patients with drug-related AEs, severe AEs, AEs with outcome of death, SAEs and adverse events leading to permanent discontinuation of treatment will be presented by study period and treatment approach.

The number and percentage of patients reporting any adverse event during the treatment period of the study will be summarized by standardized MedDRA query (SMQ) level and treatment approach as well.

In addition, the number and percentage of patients reporting any adverse event during the treatment period will also be summarized by primary system organ class, preferred term, treatment approach and age group (as defined in section 2.1.5)) overall and by treatment period..

Note that there will be no AE adjudication for this study.

2.3.1.1 Deaths, other serious and other significant adverse events

An overall summary of the number and percentage of deaths, serious adverse events (SAEs), and adverse events (including laboratory abnormalities) leading to discontinuation, study drug interruption or considered to be clinically significant will be presented by study period and treatment approach.

Deaths

The number and percentage of deaths will be summarized by study period and treatment approach and the principal cause in primary system organ class and preferred term.

Serious adverse events (SAEs)

The number and percentage of patients with serious adverse events (SAEs) will be summarized by primary system organ class, preferred term, study period and treatment approach.

Other significant adverse events – adverse events leading to discontinuations

The number and percentage of patients with adverse events that lead to discontinuation will be summarized by primary system organ class, preferred term, study period, and treatment approach.

Other significant adverse events – adverse events leading to dose adjustments or study drug interruptions

The number and percentage of patients with adverse events that lead to dose adjustments or study drug interruptions will be summarized by primary system organ class, preferred term, study period, and treatment approach.

Other significant adverse events – predefined adverse event risks

The number and percentage of patients with adverse events that were predefined risks (see [Section 1.6](#) for definitions) will be summarized by preferred term, predefined risk category, maximum severity, study period, and treatment approach.

2.3.2 Laboratory values

The details of laboratory parameters collected, SI units (note, only SI units will be reported), and normal ranges will be listed in [Section 16 of CSR](#). Note that lipid parameters are summarized along with other biochemistry laboratory parameters. Meal test laboratory parameters (glucose, insulin and c-peptide) are for use in efficacy analyses only and so should not be included in safety laboratory outputs.

Notable ranges (absolute) criteria of selected laboratory variables are defined in [Section 1.8](#) above. The laboratory values to be considered for notable ranges or relevant changes will include all post-baseline values up to the final scheduled visit, including values at unscheduled visits prior to the last visit. Only treatment emergent notable abnormalities will be considered (i.e. notably abnormal measurements which were not present at any pre-baseline visit). In addition to summary tables, a listing of patients with any treatment emergent notable laboratory results during the treatment period will be provided by treatment approach and patient number, along with their values at all visits, including repeated, unscheduled and follow-up laboratory values. In the listing, all abnormal values outside of normal range and notable values will be flagged, e.g., L/H for a value being below/above normal range, LN/HN for a value being in notable low/high range.

For each numerical laboratory parameter, the laboratory values will be reported in listings to the same precision as that for the unit in UNITCONV dataset. The precision to use in tables is as per [Section 1.3.1](#).

For continuous variables with a value <lower limit, these will be imputed as being half of the lower limit. For continuous variables with a value >upper limit, these will be imputed as being equal to the upper limit.

2.3.2.1 Hematology

The values of hematology laboratory test results at baseline and endpoints (End of Period 1, End of Period 2, End of Study) will be summarized, as well as the changes from baseline to each endpoint.

The maximum/minimum value obtained during the treatment period, and the change from baseline to maximum/minimum value will be summarized by treatment approach for each hematology parameter.

The number and percentage of patients with abnormal hematology test results meeting notable criteria post-baseline, which were not present prior to randomization, will be tabulated. This summary will be repeated for abnormalities occurring in Period 1.

2.3.2.2 Biochemistry

The values of biochemistry laboratory test results at baseline and endpoints (End of Period 1, End of Period 2, End of Study) will be summarized, as well as the changes from baseline to each endpoint.

The maximum/minimum value obtained during the treatment period, and the change from baseline to maximum/minimum value will be summarized by treatment approach for each hematology parameter.

The number and percentage of patients with abnormal biochemistry test results meeting notable criteria post-baseline, which were not present prior to randomization, will be tabulated. This summary will be repeated for abnormalities occurring in Period 1.

The number and percentage of patients with hepatic enzyme elevations (defined in [Table 1-3](#)), and persistent hepatic enzyme elevations will be presented, and the patients meeting any of the criteria will be listed. Note that for this study total bilirubin will be used for all criteria involving bilirubin. If at any visit total bilirubin has not been measured then the elevation will not be assessed for that visit (even if direct bilirubin has been measured).

The liver panel A testing data for patients with hepatic enzyme elevations will be listed at patient level. Liver panel B testing data will not be considered as part of the RAP, but will be included in narratives if applicable.

2.3.2.3 Urinalysis

The baseline and endpoints (End of Period 1, End of Period 2, End of Study) values of categorical urinalysis test results will be summarized in number and percentage by treatment approach. The baseline and endpoints (End of Period 1, End of Period 2, End of Study) values of continuous urinalysis test results will be summarized by treatment approach using descriptive statistics.

2.3.3 Vital signs

Baseline and endpoint values for blood pressure (sitting), and pulse (sitting) will be presented. Summary statistics for vital signs data will be presented by study visit. The summary by study visit will include only values obtained at the scheduled visit.

The number of patients with notably abnormal vital signs [REDACTED] post-baseline will be presented. This summary will be repeated for abnormalities occurring in Period 1. The criteria for notable abnormalities (see [Section 1.9](#)) will be listed.

2.3.4 Electrocardiograms (ECGs)

Overall interpretation and details of abnormalities are collected at Screening and selected visits.

Change from the baseline ECG result to the endpoint (End of Period 1, End of Period 2, End of Study) ECG result will be summarized by a shift table. The number of patients with clinically significant abnormalities or not at baseline and at each endpoint ECG assessment will be tabulated.

2.3.5 Hypoglycemic events

Hypoglycemic events will be characterized by event profile, such as ability to self-treat, self monitoring of plasma glucose level, precipitating event, time from last meal, time from last dose, and time of the day.

The incidence of hypoglycemic events (coded as hypoglycemia) and asymptomatic low blood glucose (coded as Blood Glucose Decreased), entered into the Glycemia study diary eCRF and meeting the event criteria defined in [Table 1-1 \(Section 1.7\)](#), are included in all of the AE summaries in [Section 2.3.1](#).

In addition, patients reporting at least one hypoglycemic event, the subgroup of reporting 1 event, 2 events, and >2 events, patients discontinued due to hypoglycemic events, patients reporting grade 2 events, and patients reporting suspected grade 2 events will be summarized by numbers and percentages for each treatment approach overall (including study periods 1, 2 and 3) and by study period. The summary will be repeated by age group (as defined in section 2.1.5).

The hypoglycemic events will further be summarized by event profile as follows:

- plasma glucose level ≤ 2.2 , $>2.2-\leq 2.8$, $>2.8-<3.1$ and ≥ 3.1 mmol/L, or Not recorded; note that if a patient is unable to initiate self-treatment and plasma glucose is not recorded, this is graded as a suspected grade 2 event)
- grade (grade 1, grade 2 or suspected grade 2)
- symptom (in preferred term)
- precipitating event (None, Missed/delayed meal, Strenuous exercise, Alcohol consumption, Other or Not recorded)
- severity (Mild, Moderate, Severe)
- relationship to study drug (Not suspected, Suspected)
- action taken (Study drug dosage adjusted / temporarily discontinued, Study drug permanently discontinued, Concomitant medication taken, Non-drug therapy given (includes oral carbohydrates), Hospitalization / prolonged hospitalization, no action taken)
- time interval between last meal and event (0–4, >4–6, >6 hours, not recorded),
- time interval between last dose and event (0–4, >4–6, >6 hours, not recorded),
- time of the day in 24-hour clock (>00:00–06:00, >06:00–12:00, >12:00–18:00, or >18:00–24:00),

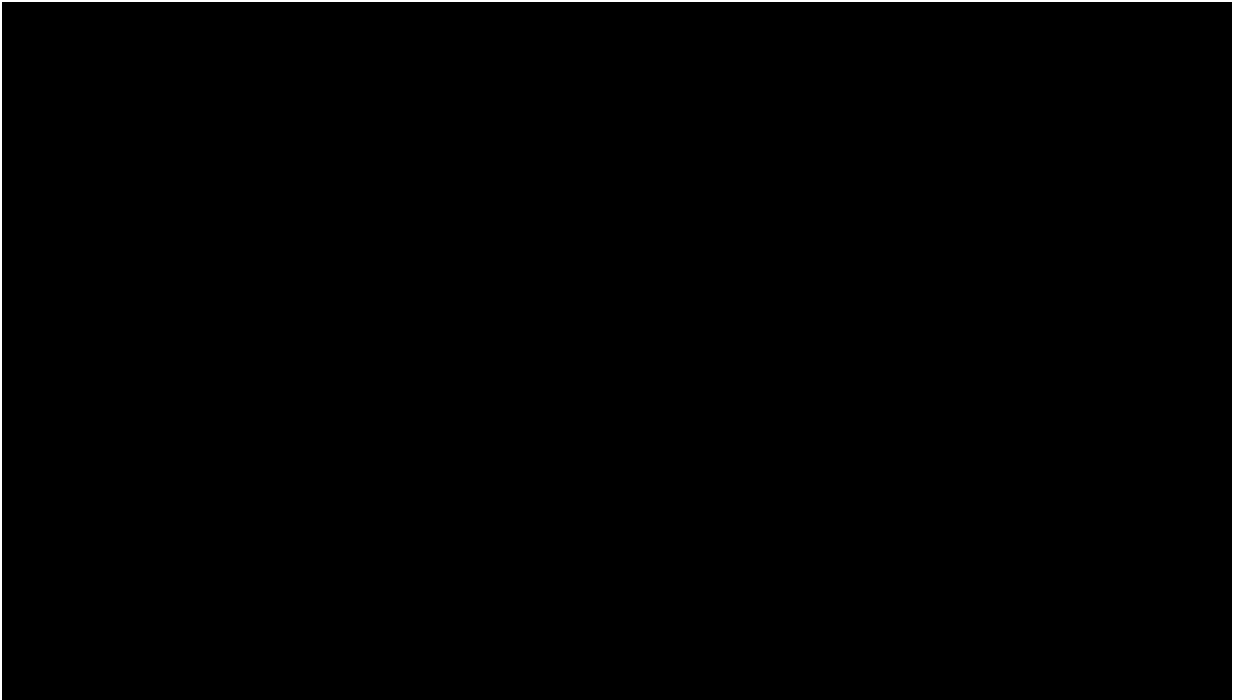
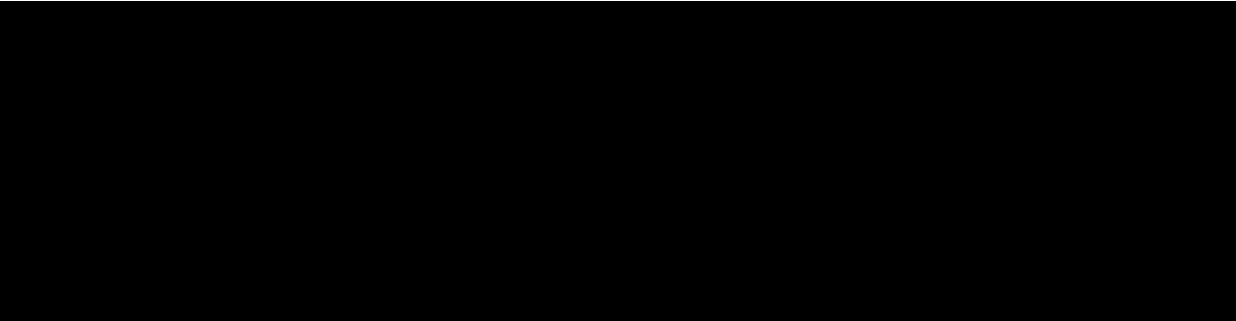
Asymptotic blood glucose events will be similarly summarized by event profile and the number of patients with at least one asymptomatic low blood glucose measurement will also be summarized.

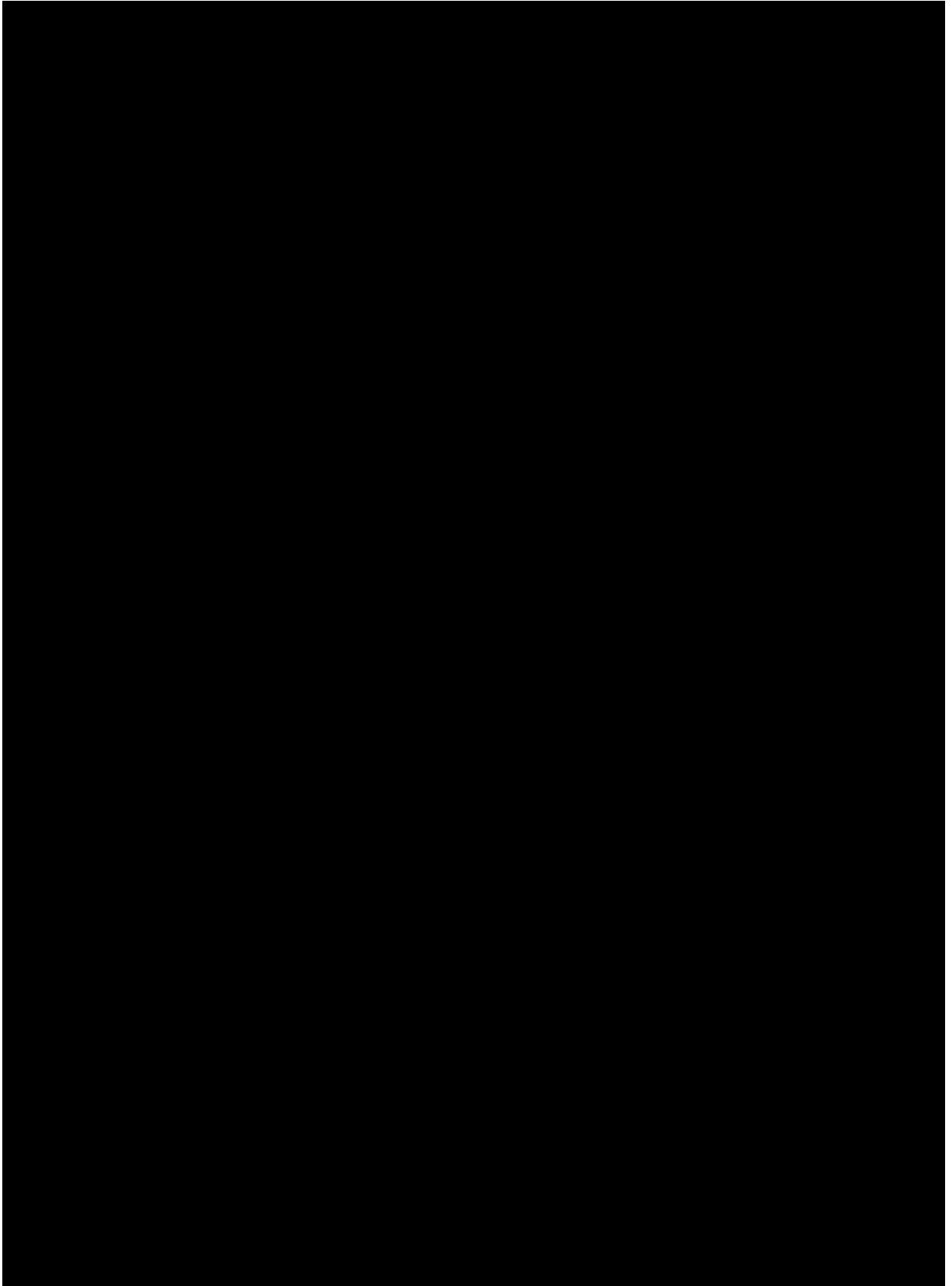
A listing of hypoglycemic events and asymptomatic low blood glucose events will be presented by treatment approach and patient number.

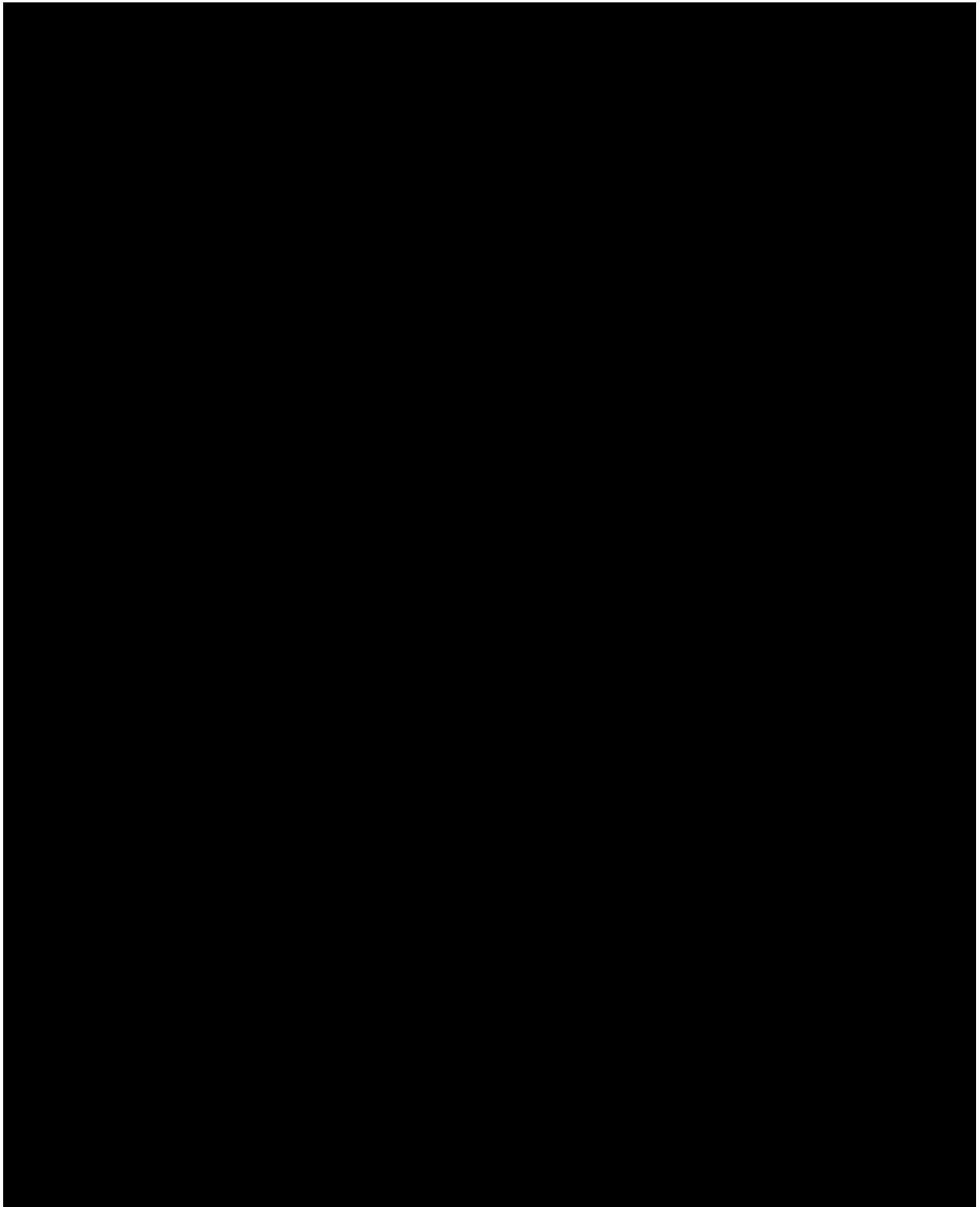
Note: As defined in [Table 1-1](#), for patients who are able to initiate self-treatment (where necessary) symptoms suggestive of hypoglycemia which are not confirmed by glucose measurements will be recorded in the Glycemia diary eCRF. Since these are not confirmed hypoglycemia events, these symptoms will not be reported in the hypoglycemia summary tables and listings, these symptoms will appear only in the AE tables and listings. For patients unable to initiate self-treatment (where necessary) symptoms suggestive of hypoglycemia that were not confirmed by glucose measurements will be classed as suspected hypoglycemic events and will therefore be reported in the hypoglycemia summary tables and listings, as well as the AE listings.

Note: Self-monitoring glucose can be recorded in the database as either plasma or whole blood. However, only plasma glucose will be reported in the tables and listings. Thus blood glucose values will be converted into plasma glucose equivalent, which can be done by multiplying a conversion factor of 1.12, i.e., plasma glucose = blood glucose \times 1.12.

The incidence of hypoglycemia and severe hypoglycemia during Period 1 will be compared for the SAF using a chi-square test between the two treatment approaches for all patients provided that the incidence such these events is more than 5 in each treatment approach.







2.5 Interim analyses

No interim analyses are planned for this study.

2.6 Other topics

No other topics will be studied.

2.7 Determination of sample size

A total sample size of 1000 randomized patients per treatment arm (in 1:1 allocation ratio to vildagliptin + metformin and metformin monotherapy) is planned.

The sample size calculation assumed that all randomized patients are to be followed up for 5 years unless patients dropped out from the study for various reasons (lack of efficacy, AEs, abnormal labs, lost to follow-up etc.). Including the updated information of the yearly dropout rate being approximately 4% (based on observed discontinuation rate for this study at the time of completion of the trial).

The existing vildagliptin study data suggest that approximately 10% of vildagliptin patients will have an HbA1c >7.0% after the first 3 months of the study (initial response phase), since those patients who are randomized with an HbA1c measurement above the failure threshold (7.0%) may never have an HbA1c measurement below the required threshold during the study and will therefore be counted as failures during the first 13 weeks. A similar proportion is assumed for the comparator arm. Hence it is expected that the difference in failure rate is likely to be small early in the study, but will diverge as the study progresses. The power calculations have been adjusted to take this assumption into account using statistical simulations.

The simulations showed that assuming an annual initial treatment failure rate of 7.1% in the metformin monotherapy arm (estimated based on ADOPT data), incorporating a 10% initial failure rate after 13 weeks in each treatment arm (due to some patients with baseline HbA1c $\geq 7.0\%$), 1000 patients per treatment arm would be sufficient to detect a hazard-ratio of 0.75 between vildagliptin + metformin and metformin alone (corresponding to a risk reduction rate of 25% in vildagliptin + metformin group versus metformin alone) with approximate 75% power and a 1-sided significance level of 0.025 (corresponding to a 2-sided test at 0.05).

2.8 Changes in the conduct of the study or planned analyses

The following change in the analysis is proposed in protocol amendment 3 (4th Oct 2016) which defers with the analysis proposed in this document. These proposed changes have come about through development of the RAP modules and discussions around the analyses.

- Treatment start date to be defined as the earliest vildagliptin/placebo date, as recorded on the DAR page of the eCRF (i.e. do not use visit 3 date). Consistently use this definition for day 1 and durations related to day 1.
- '+1' should be included in duration of exposure calculation when last study drug date is available.
- '+1' should not be included in duration of exposure calculation when last study drug date is missing/incomplete.
- Further detail added to better define the analysis sets.
- Geographic region to be included as a classification variable in the analysis of time to initial treatment failure (and all similar analyses).
- In the analysis of rate of loss in glycaemic control over time, time will be included as a fixed effect and as a random effect.
- In the analysis of time to insulin initiation, patients who complete the study in Period 1 or Period 2 will be censored at their last study visit.
- OGIS formula revised as per 23135 study.
- Due to the large number of patients (patients will be recruited until we have the required amount) and the robustness of analyses methods we can assume we will have sufficient patients data for the beta-cell function, insulin sensitivity, AUC of ISR/G and OGIS analyses.

█ [REDACTED]

█ [REDACTED]

- The definition of initial treatment failure is revised to be the date of the second of the two consecutive measurement of HbA1c value $\geq 7\%$. The rationale of this change is that this will be consistent with the Period 1 end date definition. The decision of this definition update is documented in the RAP II meeting minutes in RAP Module 1.
- Main analysis approach includes only one primary endpoint in this SAP. As the probability and risk of time to initial loss of glycaemic control as primary efficacy variable is presented as a hazard ratio and failure rate over time, the statistical approach and power calculations (including the impact of the known retention) was updated accordingly while moving the assessment of the rate of loss of glycaemic control by an annualised slope of mean HbA1c over time from week 26 to the end of period 1 to be a key secondary analysis instead.

- Baseline characteristics driving the pre-planned sub-analyses to a previously unpredictable direction, is clearly documented in the SAP to ensure transparency and external validity of the study

█ [REDACTED]

References

[ADA-EASD Consensus Statement \(2009\)](#) Medical management of hyperglycaemia in type 2 diabetes mellitus: a consensus algorithm for the initiation and adjustment of therapy. *Diabetologia*; 52:17-30.

[Hochberg and Tamhane \(1987\)](#) Multiple Comparison Procedures. Wiley Series.

[Mari A, Ludvik B, Pacini G, et al \(2001\)](#) A model-based method for assessing insulin sensitivity from the oral glucose tolerance test. *Diabetes Care*; 24: 539-548.

[Rigalleau V, Lasseur C, Perlemoine C, et al \(2005\)](#) Estimation of glomerular filtration rate in diabetic subjects. Cockcroft formula or modification of diet in renal disease study equation? *Diabetes Care* 28:838-843.

[REDACTED]

[Scherbaum W, et al. \(2008\)](#) Evidence that vildagliptin attenuates deterioration of glycaemic control during 2-year treatment of patients with type 2 diabetes and mild hyperglycaemia. *Diabetes Obes Metab*; 10:1114–1124.

[REDACTED]

[SAS/STAT User's Guide](#), version 9.2, SAS Institute, Inc., Cary, North Carolina.

Clinical Development

LAF237/Galvus

Clinical Study Protocol CLAF237A23156 / NCT01528254

A 5-year study to compare the durability of glycemic control of a combination regimen with vildagliptin & metformin versus standard-of-care monotherapy with metformin, initiated in treatment-naïve patients with type 2 diabetes mellitus

RAP Module 3 – Detailed Statistical Methodology

Author: [REDACTED], Novartis Trial Statistician
[REDACTED], [REDACTED] Trial Statistician

Document type: RAP Documentation

Document status: Final Addendum 1

Release date: 17th Dec 2019

Number of pages: 41

May not be used, divulged, published or otherwise disclosed
without the consent of Novartis

Document History – Changes compared to previous version of RAP module 3.

Version	Date	Changes
2.2	28NOV2018	<p>Section 2.1.5 - Background and demographic characteristics – baseline HbA1c categories were updated to be <7% and >=7%.</p> <p>2.2.1 - Definition of geographic region for efficacy analyses – geographic region categories were added as per Novartis decision.</p> <p>2.2.2.3 – Rate of loss in glycemc control over time – time (of HbA1c measurements, in years) added as random effect to the analysis model.</p> <p>2.3.1 - Overall experience of adverse events (AEs) – statement added that AEs will be analyzed by study period as well in addition to overall (including study events from study period 1, 2 and 3) in line with DMC analyses.</p> <p>Section 2.3.1.1 – statement added that AEs will be analyzed by study period for all subsection of AEs leading to discontinuation.</p> <p>Section 2.3.5 - Hypoglycemic events – statement added that events will be analyzed by study period as well in addition to overall in line with DMC analyses. Added new category to plasma glucose level list.</p> <p>Section 2.8 – Added statement that time will be used as both a fixed and random effect in analyses of rate of loss of glycemc control.</p> <p>Section 2.8 - Changes in the conduct of the study or planned analyses – it was added that the definition of initial treatment failure is changed to regard the date of the second of the two consecutive measurements of HbA1c values >=7%.</p>
3	06May2019	<p>Section 1.4 Changed duration derivation from weeks to Months.</p> <p>Section 1.6 Added the list of AEs of special interest</p> <p>Section 2.1.2 Added information of efficacy parameters to be considered for deviation code N02 and N03.</p> <p>Section 2.1.5 Removed definition of baseline HbA1c and FPG as it was redundant. Baseline is already defined in section 1.1</p> <p>Section 2.1.5 redefined age classification based on tertiles</p> <p>Section 2.1.6 redefined exposure categories in months instead of weeks.</p> <p>Section 2.2.1 Added table mentioning grouping for geographical regions</p> <p>Section 2.2.2 Removed loss of glycemc control as primary endpoint; provided additional detail to used confirmation visit for time to initial treatment failure; Since we have only one primary endpoint, removed multiple testing methods section</p> <p>Section 2.2.2.4 Added time to first treatment failure as supportive analysis; modified age categories for subgroup analysis; added smoking status, region and beta cell functional and insulin resistance as subgroups of interest; removed ethnicity</p> <p>Section 2.2.3 Added rate of glycemc control over time as secondary endpoint; [REDACTED]</p>


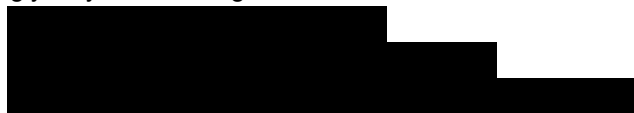
Version	Date	
		<p>analyzed and it is outside the scope of the SAP</p> <p>Section 2.2.3.3 Added the analysis of rate of glycemic control over time in this section</p> <p>Section 2.3.5 Added clarification on using chi square test to compare events only if we have sufficient observed events.</p> <p>Section 2.7 Updated power calculations based on single primary endpoint and updated yearly dropout rate.</p>
Addendum 1	17Dec2019	<p>Section 2.2.2.4 Time to second treatment failure endpoint is added as a post-hoc analysis</p> <p>Section 2.2.2.4 Subgroup analysis for the primary endpoint is added as a post-hoc analysis.</p> <p>Section 2.7 Updated sample size section based on the protocol amendment 4</p>

Table of contents

Clinical Development	1
List of abbreviations	6
1 Definitions and general strategy	7
1.1 Baseline definitions	7
1.2 Last On-Treatment Observation (Endpoint Observation)	7
1.3 General data handling strategy	8
1.3.1 Output data precision	8
1.3.2 Efficacy Assessment Window / Week X Measurement	8
1.3.3 Laboratory visit windows.....	9
1.3.4 Multiple laboratory measurements within same day	10
1.4 Duration of exposure to study medication.....	10
1.5 Treatment-Emergent Adverse Events	10
1.6 Other clinically significant events (predefined AE risks).....	10
1.7 Hypoglycemic Event Classifications	11
1.8 Laboratory abnormalities.....	12
1.8.1 Clinically Significant Laboratory Abnormalities.....	12
1.8.2 Hepatic enzyme and CPK elevations	12
1.9 Notable Vital Sign Abnormalities	13
1.10 GFR formula and definition.....	14
2 Statistical and analytical plans.....	15
2.1 Subjects and treatments	15
2.1.1 Analysis sets.....	15
2.1.2 Protocol Deviations and other criteria leading to exclusion of patients in analysis sets	15
2.1.3 Patient Disposition	16
2.1.4 Groupings for analysis	17
2.1.5 Background and demographic characteristics.....	17
2.1.6 Study medication.....	18
2.1.7 Concomitant medication	19
2.2 Efficacy evaluation	19
2.2.1 Definition of geographic region for efficacy analyses.....	19
2.2.2 Primary efficacy variables.....	20
2.2.3 Secondary efficacy variables.....	22
2.3 Safety evaluation	29
2.3.1 Overall experience of adverse events (AEs)	30

2.3.2	Laboratory values.....	31
2.3.3	Vital signs.....	33
2.3.4	Electrocardiograms (ECGs)	33
2.3.5	Hypoglycemic events	33
	[REDACTED]	34
	[REDACTED]	35
	[REDACTED]	35
	[REDACTED]	35
	[REDACTED]	36
2.5	Interim analyses	37
2.6	Other topics.....	37
2.7	Determination of sample size	38
2.8	Changes in the conduct of the study or planned analyses	39
References	41

List of abbreviations

AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATC	anatomical therapeutic chemical classification system
AUC	area under the curve
bid	bis in die (twice a day)
BMI	body mass index
BUN	Blood urea nitrogen
CPK	phosphocreatine kinase
CSR	clinical study report
ECG	electrocardiogram
eCRF	electronic case report/record form
EOS	end of study
FAS	full analysis set
FPG	fasting plasma glucose
GFR	glomerular filtration rate
HbA1c	glycosylated hemoglobin
	
ISR	insulin secretion rate
ISR/G	insulin secretion rate/glucose
LOCF	last observation carried forward
MDRD	modification of diet in renal disease
MedDRA	medical dictionary for regulatory activities
OGIS	oral glucose insulin sensitivity
PD	protocol deviation
PK	pharmacokinetics
PPS	per protocol set
RAN	randomized set
SAE	serious adverse event
SAF	safety set
SCR	screened-only set
SI	standard international
SMQ	standardized MedDRA queries
SYE	subject year exposure
T2DM	type-2 diabetes mellitus
ULN	upper limit of normal
VAS	visual analogue scale
WBC	white blood cell
WHO	world health organization

1 Definitions and general strategy

1.1 Baseline definitions

Unless specified otherwise, the baseline will be the measurement obtained on Day 1, or the closest sample obtained at an earlier visit (scheduled or unscheduled) to Day 1, if the Day 1 measurement is missing.

For derived variables [REDACTED] the baseline will be derived using measurements obtained on Day 1, or the closest samples obtained at an earlier visit (scheduled or unscheduled) to Day 1, if the Day 1 measurements are missing. Details of imputation methods to use if one or more of the individual measurements used in the derivation is missing, are detailed in [section 2.2.3.1](#).

Day 1 will be defined as the first day of study treatment (the earliest vildagliptin/placebo date as recorded on the vildagliptin/placebo dosage administration page of the eCRF).

Vital Signs

For blood pressure and pulse, baseline will be defined as the average of all pre-treatment visits, regardless of being scheduled or unscheduled. For [REDACTED], baseline will be defined as the day 1 measurement if non-missing, or last non-missing value coming from either an earlier scheduled or unscheduled visit.

Routine ECG Measurements

ECGs are analyzed and interpreted locally for this study, hence only qualitative ECG measurements are recorded. ECGs are scheduled to be measured at screening (and not Day 1). Baseline is therefore defined as the measurement obtained at screening, unless there is a later, unscheduled measurement on or before Day 1 (in which case the latest of these will be used).

[REDACTED]

[REDACTED]

1.2 Last On-Treatment Observation (Endpoint Observation)

For specified efficacy variables, clinical laboratory variables, vital signs and ECG findings, the measurement obtained at the last post-baseline study visit will be considered the last on-treatment (endpoint) observation, regardless of whether it is obtained at a scheduled or unscheduled visit, up to the last scheduled visit.

For specified efficacy variables, clinical laboratory variables and ECG findings, end of Period 1 and end of Period 2 endpoints will also be of interest. The end of Period 1 endpoint is defined as the final available post-randomization assessment obtained at any visit (scheduled or unscheduled), prior to or at initial treatment failure, up to the last scheduled study visit. The end of Period 2 endpoint is defined as the final available post-randomization assessment

obtained at any visit (scheduled or unscheduled), prior to or at the initiation of insulin therapy, up to the last scheduled study visit.

1.3 General data handling strategy

1.3.1 Output data precision

Decimal places to be used in displays of demographic, background and duration of exposure variables will be as follows:

- 3 decimal places (p-values; if p-value is less than 0.001, display <0.001),
- 2 decimal places (standard errors and standard deviations),
- 1 decimal place (means, medians),
- 1 decimal place (minimums, maximums),
- 1 decimal place (percentages),
- if percentage = 100, no decimal is required.

Decimal places for efficacy and other safety summary tables and listings will be as follows:

- 3 decimal places (p-values; if p-value is less than 0.001, display <0.001),
- data precision + 2 decimal places (standard errors and standard deviations),
- data precision + 1 decimal place (means, medians),
- same as data precision (minimums, maximums),
- 1 decimal place (percentages),
- if percentage = 100, no decimal is required.

Decimal places for the Hematology, Biochemistry and Urinalysis tables will be as follows:

- 2 decimal places for standard errors and standard deviations (Hematology and Biochemistry), 3 decimal places for standard errors and standard deviations (Urinalysis),
- 2 decimal places for means, medians (Hematology and Biochemistry), 3 decimal places for means, medians (Urinalysis),
- 2 decimal places for minimums, maximums (Hematology and Biochemistry), 3 decimal places for minimums, maximums (Urinalysis).

For vital signs tables of change from baseline, 2 decimal places for mean, median, standard deviation, minimum and maximum will be used.

In outputs containing estimates and p-values for inferential purposes (e.g. ANCOVA), means, standard errors, standard deviations and confidence intervals will be output to 2 decimal places for all parameters. P-values will be output to 3 decimal places; if a p-value is less than 0.001, it will be displayed as <0.001.

1.3.2 Efficacy Assessment Window / Week X Measurement

For all efficacy variables, the measurements obtained at the scheduled visit corresponding to *X* weeks (e.g. 13, 26, ..., 260 weeks) of treatment as defined in the assessment schedule table [Tables 6-1 and 6-2 in the protocol] will be considered as Week *X* measurement (e.g. Week 13, 26, ... , 260 measurements) if $-28 \text{ days} < (\text{visit date} - \text{Week X date}) < 28 \text{ days}$. The Week *X* date is the theoretical date when the visit should have occurred (e.g. if the scheduled visit is

visit 5, corresponding to week 26, then the Week 26 date is the treatment start date + 26 weeks).

Discontinuation visits will be assigned to the closest unallocated visit. For example, if a patient discontinues at visit 4, this discontinuation visit will be re-assigned from visit 777 to visit 4 (and not visit 23). This re-assigned visit value will be stored in a separate variable, i.e. the original value will not be overwritten.

Efficacy data obtained at visits outside the above defined window or at unscheduled visits will be excluded from the analyses and tabulation of summary statistics by visit. However, they will be displayed in the patient data listings. If the last on-treatment observations of the efficacy variables are obtained at a visit outside the above defined windows or at an unscheduled visit, they will still be used for the FAS analysis based on the last observation carried forward (LOCF) imputed endpoint (see [section 1.2](#)).

Note that prior to implementation of these windows, the remapping of the efficacy laboratory parameters will be implemented.

1.3.3 Laboratory visit windows

Safety lab data

The following will be implemented for non-efficacy laboratory data:

- Safety lab data will be summarized and listed as reported on LRS panel, no additional visit windows will be used.
- Laboratory "Visit dates" from the VIS panel on eCRF will not be listed, instead laboratory listings will present the date and day of the sample collection (obtained directly from LRS).

Efficacy laboratory data

The idea of the visit window is to ensure that as many scheduled visits as possible are included in summaries of efficacy parameters over time.

For all *efficacy laboratory summaries over time* (tables and figures), a "re-defined scheduled visit" will be created as follows:

- If scheduled visit laboratory sample already exists on database then that sample remains as the "re-defined scheduled visit".
- If scheduled visit laboratory sample is missing, then search for unscheduled sample closest to actual scheduled visit date (theoretical date visit should have occurred based on treatment start date). If unscheduled sample closest to actual scheduled visit date is within 7 days of actual scheduled date then map this unscheduled visit to the scheduled visit number in the "re-defined scheduled visit". If there is a tie between unscheduled samples before and after the actual visit, take the later measurement. Note that the randomization visit laboratory measurement will not consider measurements from unscheduled visits taken after Day 1.

All efficacy outputs will include the following footnote: *"In the case of a missing scheduled visit sample, the closest unscheduled visit within 7 days of scheduled visit is used."*

1.3.4 Multiple laboratory measurements within same day

It is possible that 2 sets of laboratory measurements may be taken with the sample date. All such measurements will be included in the summaries and listings of notably abnormal or percent change from baseline summaries. For the analysis of “persistent” events (where consecutive measurements are checked) or for the analysis of parameter over time in the following key safety parameters, only the highest within-day value will be included:

- ALT
- AST
- Bilirubin (direct and/or total)
- CPK

Specifically, in the case of duplicate measurements with same sample date, only the maximum value will be selected. All samples will be listed along with sample date.

1.4 Duration of exposure to study medication

Treatment start date is defined as the earliest vildagliptin/placebo date as recorded on the vildagliptin/placebo dosage administration page of the eCRF).

Overall duration (in months) on study medication will be computed as follows:

If a complete last study drug date is available:

$$(\text{last study drug date} - \text{treatment start date} + 1) / 30.4375,$$

or if the last study drug date is missing or incomplete:

$$(\text{last visit date} - \text{treatment start date}) / 30.4375,$$

or if a patient received no dose of randomized double-blind medication:

overall duration is set to 0.

The study drug end date is defined as the last treatment date. The last treatment date in the study completion panel will be used for this study drug end date.

1.5 Treatment-Emergent Adverse Events

A treatment-emerging adverse event is defined as

- an undesirable sign, symptom, or medical condition with onset after the first dose of study medication or,
- an increase in severity of an event that is present during the pre-treatment period,
- all events included on the database will be included in AE summary tables/listings.

1.6 Other clinically significant events (predefined AE risks)

For this project, there are some predefined adverse event risks that are considered of special interest defined as below:

Important identified risks

- Transaminase elevations and Drug-induced liver injury (DILI)
- Angioedema

- Acute pancreatitis
- Skin lesions (pemphigoids)
- Hypoglycemia
- Lactic acidosis (metformin)

Important potential risks

- Serious infections
- Cardiac events in CHF (NYHA Functional Class III) patients (adjudication)
- Muscle events/ myopathy/rhabdomyolysis
- Neuropsychiatric events
- Breast cancer
- Pancreatic cancer

1.7 Hypoglycemic Event Classifications

Patients were educated to record hypoglycemic symptoms and treatment in a study diary. Study diary data were then entered in the Glycemia Study Diary or Adverse Event eCRF, respectively according to the following criteria. The Glycemia study diaries were reviewed by the study staff at each visit and assessed according to [Table 1-1](#). Note that if glucose is recorded as whole blood on the eCRF, rather than plasma, then it will be converted to a plasma value in the derived dataset, using a conversion factor of 1.12 (i.e. plasma glucose = blood glucose x 1.12).

Table 1-1 Recording and classification of glycemia study diary data

Severity	Symptoms	Plasma glucose	Action taken	Classification
Patient able to initiate self-treatment if necessary	Symptoms suggestive of hypoglycemia	< 3.1mmol/L (56mg/dL)	Enter in Glycemia Study Diary eCRF	Hypoglycemic event, grade 1
	Symptoms suggestive of hypoglycemia	≥ 3.1mmol/L(56mg /dL)	Enter in Glycemia Study Diary eCRF	Adverse Event
	Symptoms suggestive of hypoglycemia	Not taken	Enter in Glycemia Study Diary eCRF	Adverse Event
	Asymptomatic	< 3.1mmol/L(56mg /dL)	Enter in Glycemia Study Diary eCRF	Asymptomatic low blood glucose
Patient is unable to initiate self-treatment and requires assistance of another person or hospitalization	Symptoms suggestive of hypoglycemia	< 3.1mmol/L(56mg /dL)	Enter in Glycemia Study Diary eCRF & report as SAE	Hypoglycemic event, grade 2 & SAE
	Symptoms suggestive of hypoglycemia	Not taken	Enter in Glycemia Study Diary eCRF & report as SAE	Suspected hypoglycemic event, grade 2 & SAE

1.8 Laboratory abnormalities

Laboratory data will be presented in SI units only.

1.8.1 Clinically Significant Laboratory Abnormalities

Notable ranges for selected laboratory parameters of interest as listed in [Table 1-2](#) are the basis for the central laboratory to generate notable reports to alert the investigators. The notable ranges will form the basis to evaluate the incidence of clinically significant laboratory abnormalities.

Table 1-2 Notable ranges (absolute) criteria of laboratory variables of specific interest

Lab test description	SI Unit	Notable ranges	
		Low	High
Hematology			
Total leucocytes (WBC)	10E9/L	≤ 2.8	
Total leucocytes (WBC)	10E9/L		≥ 16.0
Platelet count (direct)	10E9/L	≤ 75	
Platelet count (direct)	10E9/L		≥ 700
Hemoglobin	g/L	≤115 (m) ≤ 95 (f)	–
Hematocrit	1	<0.37 (m) ≤ 0.32 (f)	–
Eosinophils	%	–	≥ 14
Biochemistry			
Sodium	mmol/L	≤ 125	
Sodium	mmol/L		≥ 160
Potassium	mmol/L	≤ 3	
Potassium	mmol/L		≥ 6
Blood urea nitrogen (BUN)	mmol/L	–	≥ 14.28
Creatinine	μmol/L	–	≥ 176.8

1.8.2 Hepatic enzyme and CPK elevations

In addition to the notable ranges, the incidence of treatment emergent hepatic enzyme elevations is of interest. These are described in [Table 1-3](#).

Treatment emergent is defined as not meeting the criterion at any pre-baseline visit, but then meeting the criterion at a post-baseline visit. Consequently a patient may be classified treatment emergent ≥5ULN even though they were not classified as treatment emergent ≥3ULN (if had a result ≥3ULN at a pre-baseline visit).

Table 1-3 Treatment emergent hepatic enzyme and CPK elevations of interest

ALT or AST ≥ 3*ULN

ALT or AST \geq 5*ULN
ALT or AST \geq 8*ULN
ALT or AST \geq 10*ULN
ALT or AST \geq 20*ULN
ALT or AST \geq 3*ULN and Bilirubin[^] \geq 1.5*ULN
ALT or AST \geq 3*ULN and Bilirubin[^] \geq 2*ULN
Bilirubin[^] \geq 2*ULN
Bilirubin[^] \geq 3*ULN
CPK \geq 5*ULN
CPK \geq 10*ULN
ALT \geq 3*ULN
ALT \geq 5*ULN
ALT \geq 8 x ULN
ALT \geq 10*ULN
ALT \geq 20*ULN
ALT \geq 3*ULN and Bilirubin[^] \geq 1.5*ULN
ALT \geq 3*ULN and Bilirubin[^] \geq 2*ULN
AST \geq 3*ULN
AST \geq 5*ULN
AST \geq 8 x ULN
AST \geq 10*ULN
AST \geq 20*ULN
Alkaline Phosphatase \geq 1.5*ULN

[^] If only direct bilirubin is measured then direct bilirubin will be used for all criteria involving bilirubin. If both total and direct bilirubin are measured (i.e. are in the laboratory specifications) throughout a study, total bilirubin will be preferred. The clinical study team has agreed that total bilirubin should be used for this study. Note that if at any visit total bilirubin has not been measured then the elevation will not be assessed for that visit (even if direct bilirubin has been measured).

In addition to the overall summary of elevations defined above, the same criteria will also be presented, restricted to “persistent elevations” only. Persistent elevations are those which meet the criterion at consecutive on-treatment measurements or at last on-treatment visit (this includes assessments within one day of the last dose of study treatment). The purpose of the persistent definition is to remove transient elevations from the summary (i.e. those elevations which appeared at one visit but normalized at the next visit).

1.9 Notable Vital Sign Abnormalities

Notable vital sign abnormalities [REDACTED] are defined in [Table 1-4](#).

Table 1-4 Notable vital sign abnormalities

Vital signs	Notable abnormalities		
	Low	High	
Pulse (beats/min)	either ≤ 50 + decrease $\geq 30^*$ or < 40	either ≥ 120 + increase $\geq 25^*$ or > 130	
Blood pressure (mmHg)	Systolic	either ≤ 90 + decrease $\geq 30^*$ or < 75	either ≥ 180 + increase $\geq 30^*$ or > 200
	Diastolic	either ≤ 50 + decrease $\geq 20^*$ or < 40	either ≥ 105 + increase $\geq 20^*$ or > 115
██████████	██	██	

* Refers to post-baseline value as compared to baseline value

The notable abnormality of ‘ ≥ 120 + increase ≥ 25 or > 130 ’ will be interpreted as that the value in the treatment period is ≥ 120 and the increase from baseline is ≥ 25 , or the value in the treatment period is > 130 .

1.10 GFR formula and definition

The following formula is used to estimate Glomerular Filtration Rate (GFR) modification of diet in renal disease (MDRD) (Rigalleau 2005):

$$\begin{aligned}
 \text{GFR}_{\text{MDRD}} (\text{mL/min}) \times (1.73/\text{m}^2) = & \\
 & 175 \times (\text{serum creatinine } \{\text{mg/dL}\})^{-1.154} \\
 & \times (\text{age } \{\text{years}\})^{-0.203} \\
 & \times 0.742 \text{ (if female)} \\
 & \times 1.210 \text{ (if race is Black)}
 \end{aligned}$$

GFR will be available in the raw laboratory data (and used to assess inclusion/exclusion criteria) but will be re-derived within the derived datasets using this RAP-specified formula. This will ensure consistency with other studies within the program.

Baseline GFR (MDRD) is a derived variable calculated using the baseline serum creatinine, ██████████ & age value, i.e. the value obtained Day 1, or the sample obtained at an earlier visit (scheduled or unscheduled) which is closest to Day 1, if the Day 1 measurement is missing. Hepatic function is categorized as follows:

- Normal: GFR (MDRD) > 80
- Mild: GFR (MDRD) ≥ 50 - GFR (MDRD) ≤ 80
- Moderate: GFR (MDRD) ≥ 30 - GFR (MDRD) < 50
- Severe: GFR (MDRD) < 30

2 Statistical and analytical plans

Data will be analyzed [REDACTED] according to the data analysis [section 9 of the study protocol](#) which is available in [Appendix 16.1.1 of the CSR](#). Important information is given in the following sections and details are provided, as applicable, in [Appendix 16.1.9 of the CSR](#).

2.1 Subjects and treatments

2.1.1 Analysis sets

Screened-only set (SCR): The SCR consists of all patients who were screen failed after the first visit or who entered the run-in phase but were not randomized. Except for the listing of individual patients in this SCR with the reasons for not being randomized, and the tabulation of patients in this SCR by discontinuation reason, no other analysis will be performed on this analysis set.

Randomized set (RAN): The RAN consists of all randomized patients.

Full analysis set (FAS): The FAS consists of all randomized patients who receive at least one dose of randomized study medication (vildagliptin or placebo) and have at least one post-randomization assessment of any efficacy parameter. Following the intent-to-treat principle, patients will be analyzed according to the treatment approach they are assigned to at randomization.

Safety set (SAF): The SAF consists of all patients who receive at least one dose of randomized study medication (vildagliptin or placebo). Patients will be analyzed according to the treatment approach received. If a patient receives both vildagliptin and placebo in Period 1, then the patient will be included in the vildagliptin group. Note that the SAF allows the inclusion of non-randomized patients who received the study drug in error.

Per Protocol set (PPS): The PPS is a subset of FAS and consists of all randomized patients who receive at least one dose of randomized study medication (vildagliptin or placebo), have at least one post-randomization assessment of any efficacy parameter in Period 1, do not discontinue the study prior to Week 26, and have no major protocol deviations occurring during Period 1.

2.1.2 Protocol Deviations and other criteria leading to exclusion of patients in analysis sets

Patients will be excluded from analysis based on the protocol deviations (PDs) in [VAP Module 3](#) and other criteria described in [Table 2-1](#). The PDs described in [VAP Module 3](#) will be identified manually by field monitors/CTH and/or via PD programming, and will be provided in the VIOPTO source dataset, whilst those in ([Table 2-1](#)) will be programmed by Biostatistics using source data, and stored in the derived dataset NOVIOPTO. The severity codes assigned to each deviation as described in [VAP Module 3](#) and [Table 2-1](#) will be used to classify patients as described in [Table 2-2](#). Note that other protocol deviations will also be identified. However, patients will not be excluded from analysis based on these other deviations (with severity code 49) including violations of entry criteria on prior use of anti-

diabetic agents and BMI range. The subject classification as specified in [Table 2-2](#), will be defined based on severity codes prior to un-blinding the database with respect to treatment.

Table 2-1 Other criteria leading to exclusion of patients in analysis sets

Description of Deviation	Deviation code	Severity code	Inclusion/Exclusion in Analyses
No dose of randomized, double-blind medication (vildagliptin or placebo) taken during Period 1	N01	8	Excluded from all efficacy and safety analyses
No post-randomization efficacy assessment during the treatment period (Periods 1, 2 or 3)	N02	0	Excluded from efficacy analyses
No post-randomization efficacy assessment during Period 1	N03	1	Excluded from PPS analyses
Patient discontinued from the study prior to Week 26	N04	1	Excluded from PPS analyses

Subjects with no post-randomization efficacy assessment of Hba1c, fasting insulin, fasting glucose, fasting C-peptide and meal-test parameters, will be excluded from efficacy analyses (Deviation code N02). Subjects with no post-randomization efficacy assessment of Hba1c, will be excluded from PPS analyses (Deviation code N03).

Table 2-2 Subject classification

Analysis set	Deviation Severity Codes			
	0	1	8	49
Randomized Set (RAN)	Yes	Yes	Yes	Yes
Full Analysis Set (FAS)	No	Yes	No	Yes
Safety Set (SAF)	Yes	Yes	No	Yes
Per Protocol Set (PPS)	No	No	No	Yes

Note: yes/no means whether patients with PDs or non-PD criteria of a specific code are included (yes) or excluded (no) from an analysis set.

The number and percentage of patients in the RAN with protocol deviations and other criteria leading to exclusion of patients from analysis sets will be tabulated by treatment approach. Protocol deviations that do not lead to exclusion from analysis populations will be further summarized for the RAN by treatment approach. Summaries will be repeated by Period.

2.1.3 Patient Disposition

The number and percentage of randomized patients (RAN) who completed the study and who discontinued will be tabulated by treatment approach. Discontinuation will be further broken down by reason in the tabulation. Summaries will be repeated by Period, including information about the number of patients continuing into the next Period.

Completion information will be listed for the RAN.

Patients who were screened for the trial and entered the run-in phase but were not randomized (SCR) will be tabulated and listed with reasons for not being randomized. These summaries will include both screen failures and run-in failures (i.e. patients who discontinue within the screening period and patients who discontinue within the run-in period). Note that the patients will be presented together and not separated out into separate groups for screen failures and run-in failures.

2.1.4 Groupings for analysis

The number and percentage of patients in each analysis set will be summarized by treatment approach. The number and percentage of patients included in the meal test, [REDACTED] subsets will also be summarized by treatment approach. Meal test [REDACTED] count data will be presented as part of the efficacy evaluations. [REDACTED]

2.1.5 Background and demographic characteristics

Demographic, background data and key efficacy variables at baseline will be summarized for all randomized patients (RAN) by the treatment approach that patients are assigned to at randomization. Categorical variables gender, race, ethnicity, baseline age group, current smoking status, baseline GFR (MDRD) category, baseline body mass index (BMI) group, and baseline HbA1c group will be summarized by frequency and percentage. The continuous variables baseline age, height, [REDACTED], baseline BMI, duration of T2DM, baseline HbA1c and baseline FPG will be summarized by mean, standard deviation, median, minimum and maximum.

All baseline and background characteristics (except HbA1c, FPG and GFR) will be summarized at study entry (screening visit) or at the randomization visit if screening (Week - 5) measurements are not collected.

Baseline body mass index (BMI) will be calculated (i.e. not taken from the eCRF) from the height and weight measured at screening (Week -5), $BMI = \text{weight} / \text{height}^2$, where weight is in kilograms (kg) and height is in meters (m).

The classification of the baseline HbA1c, BMI, age category and GFR (MDRD) category are based on the following definitions:

- HbA1c category (< 7% and $\geq 7\%$)
- BMI (< 30 kg/m² and ≥ 30 kg/m² at Visit 1)
- Age at Visit 1 by tertiles i.e. <48 years, 48 - < 62 years and ≥ 62 years)
- GFR (MDRD) category (Normal: >80 mL/min, Mild: ≥ 50 mL/min - ≤ 80 mL/min, Moderate: ≥ 30 mL/min - <50 mL/min, Severe: 0 mL/min - <30 mL/min)

Baseline comparability across treatment approaches for all randomized patients (RAN) will be examined in a [Section 16.1 listing](#) using a chi-square test for the categorical variables, and a two-sample t-test for the continuous variables. The p-values are provided for descriptive purposes, and will not to be considered as the formal basis for determining factors to be included in the statistical analysis model.

In addition, all relevant medical history and history of diabetes and complications will be summarized by primary system organ class, preferred term and treatment approach using frequency tables in Post-text tables for all RAN patients.

2.1.6 Study medication

The study consists of three treatment periods.

In Period 1 patients will be randomized to metformin + vildagliptin 50mg bid, or metformin + placebo. If initial treatment fails to maintain HbA1c < 7.0%, confirmed at two consecutive scheduled study visits (after remapping in [section 1.3.3](#) has been applied), starting from Visit 4 (Week 13), then vildagliptin will be added to metformin in the metformin-only arm in Period 2. The Period 1 end date is defined as the date of the second of the two consecutive scheduled study visits, i.e. the initial treatment failure date. The Period 2 start date is defined as the next day.

In Period 3, insulin is considered based on local guidelines. Both patients and investigators will remain masked to the treatment allocation in Period 1, and the study will compare the two different treatment approaches. The Period 3 start date is defined as the date of insulin initiation. The Period 2 end date is defined as the previous day.

Treatments are referred to as ‘Treatment approach’ throughout this analysis plan to acknowledge that patients are not on the same treatment for the duration of the study. When summaries are detailed for the ‘Treatment Period’, they will include data from Periods 1, 2 and 3, unless specified otherwise.

The duration of exposure to study medication (in months) during the treatment period will be computed for all patients using the algorithm outlined in [Section 1.4](#) and summarized by treatment approach for the RAN both descriptively (i.e., mean, standard deviation, median, minimum and maximum) and by 3 - monthly intervals (0 - < 3 months, 3 - < 6 months, ..., ≥ 60 months).

The study Period 1 duration (in months) will be also similarly summarized for the RAN. The study Period 1 duration is defined as (initial treatment failure date – treatment start date + 1)/30.4375 for patients who entered study Period 2. The study Period 1 duration is the same as overall study duration for patients who did not enter into Period 2.

The study period from treatment start to the end of Period 2 will be computed for all patients who initiate insulin in Period 3 and for patients discontinuing the study in Period 2 due to being unable or unwilling to initiate insulin for treatment intensification in Period 3. The duration will be defined as (initiation date of insulin – treatment start date)/30.4375 for patients who initiate insulin, and (last study drug date – treatment start date + 1)/30.4375 for patients who do not enter Period 3.

Total subject year exposure (SYE) will also be summarized by treatment approach for the RAN. SYE (years) for each treatment will be calculated as (duration of exposure in days, disregarding any treatment interruptions, for all subjects in that treatment approach) / 365.25. For patients who did not take any study medication, the exposure is set to zero. This summary will be repeated on Period 1 data only, and also on Vildagliptin exposure only.

Study medication information will be listed for the RAN, including information about vildagliptin, placebo, metformin and insulin use.

2.1.7 Concomitant medication

Concomitant therapies taken during the treatment period will be summarized by anatomical therapeutic chemical classification system (ATC) class and preferred term according to the treatment approach patients are assigned to at randomization in the RAN. The number and percentage of randomized patients (RAN) in each class and preferred term will be tabulated by treatment approach. This summary will be repeated on period 1 data only, i.e. concomitant therapies taken during Period 1.

Use of anti-diabetic medications prior to study entry (as recorded on the eCRF) will be summarized by ATC class, preferred term and treatment approach.

The duration and average daily dose of metformin taken during the screening and run-in period will be summarized descriptively (mean, standard deviation, median, minimum and maximum), using data from 35 days prior to screening through to 35 days after randomization. This will be based on information collected on the i) Metformin dosage administration page of the eCRF (considering only dosages which are at most 35 days after randomization); and ii) information collected on the prior anti-diabetic medication eCRF page with a preferred term containing 'METFORMIN' and a dose of 'mg' (considering only dosages which are at most 35 days prior to the screening visit). The average daily dose of metformin during this window will also be summarized categorically (<2000 mg, ≥2000 mg).

Listings of concomitant medications and prior anti-diabetic medications will be presented by treatment approach and patient number.

2.2 Efficacy evaluation

2.2.1 Definition of geographic region for efficacy analyses

Centers/countries with a small number of patients will be pooled to create geographic regions. Larger centers/countries will be considered as their own geographic region. There are six geographic regions planned: Europe, Latin America, South-East Asia, East Asia, Australia and Africa. The pooling of centers and/or countries to create geographic regions will be confirmed once all patients have been randomized. The allocation of countries to the specific geographic regions will be as follows.

Centers/Countries	Geographic region
AUT, BGR, CZE, DEU, ESP, EST, FIN, HUN, ISR, ITA, LTU, LVA, NOR, POL, ROU, RUS, SVK, TUR	Europe
ARG, BRA, COL, DOM, GTM, MEX, PAN, PER	Latin America
IND, MYS, PHL	South-East Asia
HKG, KOR, TWN	East Asia

AUS	Australia
ZAF	Africa

2.2.2 Primary efficacy variables

2.2.2.1 Variable definition

The primary efficacy variable is time to confirmed initial treatment failure. The confirmed initial treatment failure is defined as a patient with HbA1c measurements $\geq 7.0\%$ at two consecutive scheduled visits (after remapping in [section 1.3.3](#) has been applied), starting from Visit 4 (Week 13). The time to confirmed initial treatment failure is defined as the time from randomization until the time when the second of the two HbA1c measurement $\geq 7.0\%$ was determined after at least 13 weeks of treatment (Visit 4), i.e. the end of Period 1. The rationale of starting the treatment failures count from Visit 4 is that treatment guidelines consider as treatment failure patients who do not achieve HbA1c $< 7.0\%$ after 3 months of treatment ([ADA-EASD, 2009](#)).

The main study conclusion will be based on the analysis of the primary efficacy variable using the FAS. Analyses based on the PPS will also be performed to assess the robustness of the conclusion.

2.2.2.2 Handling of missing values/censoring/discontinuations

In the analysis for the primary efficacy variable of time to treatment failure, patients who discontinue the study for any reason during Period 1 (lack of efficacy, lost to follow-up, AE or abnormal laboratory values etc.) will be treated as censored values at the time of discontinuation. Patients who remain under the threshold (or whose measurement above the threshold is not confirmed at next scheduled visit) will be censored at the time of last study visit.

No imputation will be used for missing HbA1c measurements. As detailed in [section 1.3.3](#), re-defined scheduled visits will be created and so it should be the case that there is always a measurement for a scheduled visit, as an unscheduled visit should be carried out to ensure there was an HbA1c value available. However, if there is a case where the HbA1c value is missing then when looking for consecutive scheduled visits, only visits with non-missing HbA1c measurements will be considered (e.g. if Visit 4 HbA1c $\geq 7.0\%$, Visit 5 HbA1c is missing, and Visit 6 HbA1c $\geq 7.0\%$, then this will be considered treatment failure).

2.2.2.3 Statistical model, hypothesis, and method of analysis

Time to initial treatment failure

The primary statistical hypothesis of time to initial treatment failure will be assessed by a 1-sided test of superiority of the combination treatment with vildagliptin + metformin versus metformin monotherapy with α -level of 0.025 with a null hypothesis that the hazard-ratio between the combination treatment vildagliptin + metformin and metformin monotherapy is equal or greater than 1, versus the 1-sided alternative hypothesis that the hazard-ratio is less than 1:

$H_0: \lambda_{\text{vildagliptin+metformin}}/\lambda_{\text{metformin}} \geq 1$

versus

$H_a: \lambda_{\text{vildagliptin+metformin}}/\lambda_{\text{metformin}} < 1$

where $\lambda_{\text{vildagliptin+metformin}}$ and $\lambda_{\text{metformin}}$ are the hazard-rates of the initial treatment failure for vildagliptin 50mg bid + metformin up to 1000mg bid and metformin up to 1000mg bid monotherapy respectively.

The time to confirmed failure will be derived as the time from randomization to the second of two consecutive scheduled visits (after remapping in [section 1.3.3](#) has been applied), at which HbA1c $\geq 7.0\%$ is measured, starting from Visit 4 (13 weeks after randomization).

The primary efficacy analysis will use a Cox proportional hazard regression model to assess the probability of initial treatment failure, with treatment approach and geographic region as classification variables and baseline HbA1c as a covariate. The hazard-ratio and associated 95% confidence interval as well as the p-value estimated from the above model will be presented by treatment approach.

The initial treatment failure rate over time by treatment approach will be summarized using estimates and 95% confidence intervals from a Kaplan-Meier analysis. The Kaplan-Meier estimates will be plotted, with symbols to indicate censored values.

The primary analysis for this primary efficacy variable will be performed using the FAS. It will be performed in the PPS as supportive analysis as well.

The model assumptions will be checked by plotting the Schoenfeld residuals for the treatment effect obtained from the model against time. A random spread of residuals with no clear pattern will indicate that the proportional hazards assumption is valid. The results will be presented in CSR Section 16 table(s).

2.2.2.4 Supportive analyses

Time to first treatment failure

The time to first treatment failure will be derived as the time from randomization to the first of two consecutive scheduled visits (after remapping in [section 1.3.3](#) has been applied), at which HbA1c $\geq 7.0\%$ is measured, starting from Visit 4 (13 weeks after randomization).

Time to first treatment failure will be analyzed in the similar way as primary analysis.

Time to second treatment failure

The time to second treatment failure will be derived as the time from randomization to the second of the two consecutive scheduled visits (after remapping in [section 1.3.3](#) has been applied), at which HbA1c $\geq 7.0\%$ is measured in period 2.

Patients who discontinue the study for any reason during Period 1 or period 2 without the treatment failure are censored at the date of discontinuation. Patients that failed (2 consecutive scheduled visits with $\geq 7\%$) in period 1 and didn't failed in period 2 are at the end of period 2. Patients who didn't failed in period 1 passed to period 3 are censored to last study visit prior period 3.

Time to the second treatment failure will be analyzed in the similar way as primary analysis.

Descriptive analysis in HbA1c over time

Summaries of absolute values and change from baseline to end of Period 1 in HbA1c by treatment approach and visit will be presented in the FAS. Figures will be produced showing the mean HbA1c and standard error by visit during the study Period 1 for each treatment approach. The summary table and figure will be repeated for PPS.

In addition, the above descriptive analysis (including summary table and figure) will be performed for FAS on HbA1c data collected up to the end of Period 2, as well as data collected during the entire duration of the study.

Subgroup analysis

Summaries of absolute values and changes from baseline to end of Period 1 in HbA1c will be presented by treatment approach. The analysis will be performed on FAS. The list of subgroups of interest is listed below:

1. Baseline HbA1c category (<7%, ≥7%).
2. BMI (<30kg/m², ≥30kg/m² at Visit 1).
3. Age at Visit 1 by tertiles i.e. < 48 years, 48-< 62 years, ≥ 62 years.
4. Gender.
5. Race.
6. Smoking status
7. Geographical regions

In addition to above mentioned subgroups, following will be analyzed and reported for time to confirmed treatment failure endpoint in the same way as defined for the primary analysis.

- Age (three subgroups per tertiles of baseline age in all patients pooling the two treatment groups) (i.e., ≤ P33 years, >P33 - ≤ P66 years and > P66 years)
- Age (< 65 years, ≥ 65 years)
- Baseline GFR [mL/min/1.73m²] (<90, ≥90), baseline GFR (<60, ≥60), baseline GFR (<30, 30-<60, 60-<90).

Other baseline demographic subgroups may be added as appropriate.

2.2.3 Secondary efficacy variables

2.2.3.1 Derivation of secondary [REDACTED] variables

Rate of glycemic control over time

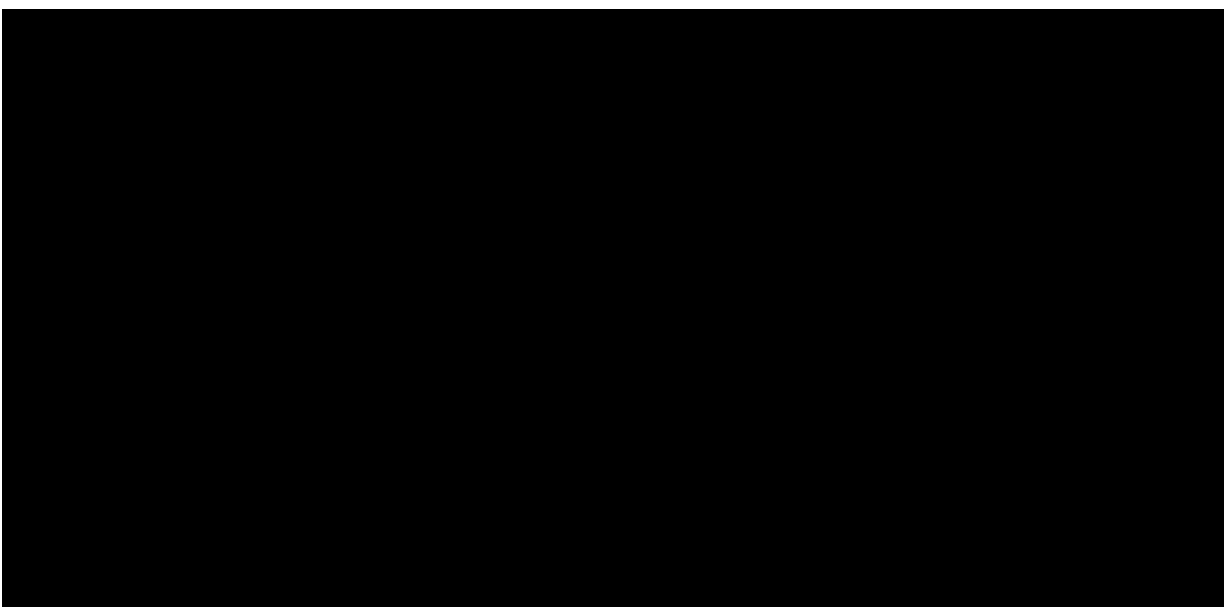
The rate of loss of glycemic control over time will be estimated by an annualized slope of HbA1c over time from Visit 5 (Week 26) to the end of Period 1. A mixed effect model

including a random coefficient will be used for analysis. The rationale of choosing Visit 5 as starting point for estimating the rate (slope) of loss of glycemetic control over time is based on data from study CLAF237A2307 (Scherbaum W et al, 2008) where the maximal reduction in mean HbA1c from baseline was between 24 and 32 weeks, therefore a trend for loss in glycemetic control is sensible to be estimated from a similar time point.

In the analysis of the rate of loss in glycemetic control over time, data used in analysis will be censored at the start of Period 2, i.e. the only data included in the model will be from Visit 5 (Week 26) to the end of Period 1.

FPG

If the FPG value (standard non meal-test parameter) is missing, the -20 min meal-test value (obtained on the same day) can be used to impute. If the -20 min value is also missing, the 0 min value will be used.



Calculation of OGIS

OGIS will be calculated as a function of glucose and insulin (meal-test laboratory parameters):

$$\text{OGIS} = 0.5 * \{ \\ \quad [(P_5 * (G(90) - G_c)) + 1] * CI_OGTT] \\ \quad + \text{sqrt} \{ \\ \quad \quad [(P_5 * (G(90) - G_c)) + 1] * CI_OGTT] ^2 \\ \quad \quad + [4 * P_5 * P_6 * (G(90) - G_c) * CI_OGTT] \\ \quad \quad \} \\ \quad \}$$

where,

- $CI_OGTT = \{ P_4 * \{ (P_1 * D_0) - [V * (G(120) - G(90))] / 30 \} / G(90) + (P_3 / G(0)) \}$
 $/ \{ I(90) - I(0) + P_2 \}$
- D_0 (g/m²) is the glucose dose included in the standard meal administered to patients who participate in the standard meal-test ($D_0 = (73.2 * 5.551) / BSA$ for the standard meal as planned in [Protocol Section 6.4.3](#). It will be assumed all patients receive a standard meal),
- $BSA = 0.164 * (\text{Weight in kg} ** 0.515) * ((0.01 * \text{Height in cm}) ** 0.422)$,
- $V = 10000$ (ml/m²) is the assumed glucose distribution volume,
- $G(0)$, $G(90)$ and $G(120)$ are glucose concentration values (taken from meal-test laboratory data) at timepoints 0, 90 and 120 min,
- $I(0)$ and $I(90)$ are insulin concentration values (taken from meal-test laboratory data) at timepoints 0 and 90 min,
- $P_1=6.5$, $P_2=1951$, $P_3=4514$, $P_4=792$, $P_5=0.0118$, $P_6=173$, $G_c=5$ are constants ([Mari, et al 2001](#)).

If the value of glucose or insulin values at 0 minutes are missing, then the -20 min meal-test value (obtained on the same day) can be used to impute. If the -20 min value is also missing, the fasting parameter result (i.e. standard non meal-test parameter) will be used. Both glucose and insulin should be imputed using the same timepoint (regardless if one has value available).

No imputation will be used for missing 90 or 120 minute values.

Calculation of ISR

Separately for each patient and each meal-test, deconvolution is applied to estimating insulin secretion to assess β -cell function. C-peptide, which is co-secreted with insulin in an equimolar ratio, is preferred to insulin as a basis for estimating insulin secretion.

Mathematically, we need to solve for the insulin secretion rate r given the impulse response h , and plasma C-peptide c , where

$$c(t) = \int_{-\infty}^t h(t-s)r(s)ds.$$

The impulse response is the PK model for an intravenously administered unit bolus dose of C-peptide when the endogenous C-peptide production has been suppressed, i.e.

$$h(t-s) = (1/V)[f e^{-\alpha(t-s)} + (1-f)e^{-\beta(t-s)}],$$

where V , f , α , and β are a subject's parameters, which are generally unknown. However, from clinical studies these parameters have been estimated, and are related to a subject's disease status, body surface area, age, and gender. Thus, h becomes known.

Insulin secretion rates are modeled by a piecewise continuous linear function with knots at the actual sampling time. Changes in slopes are penalized through a single regularization parameter that is determined by a root finding method to produce similar residual standard deviation for fitted C-peptide concentrations as would be expected from assay error standard deviations given in the kit instructions.

The IMMULITE 2000 C-peptide kit instructions were used to model intra-assay SD as a function of the true C-peptide concentration $C1$ expressed in conventional units of ng/mL by simple linear regression as:

$$SD = 0.04056C1 + 0.06311\text{ng/mL.}$$

With concentration of C-peptide expressed alternatively in SI units, i.e., $C2$ expressed in nM (1ng/mL C-peptide = 0.331nM C-peptide) the equations for SD becomes:

$$SD = 0.04056 C2 + 0.331 \times 0.06311\text{nM.}$$

The following ln-transformation with a shift parameter of observed and fitted C-peptide values is used to stabilize the variance in performing the deconvolution:

$$\ln(C1 + 0.06311/0.04056) \text{ or } \ln(C2 + 0.331 \times 0.06311/0.04056).$$

Nonlinear least squares with ISR constrained non-negative using PROC IML function NLPLM in SAS® is used to solve for ISR at the sampling time points using actual sampling times. The bisection method is used to iteratively determine the regularization parameter such that the sum of squares of the observed C-peptide minus the fitted C-peptide expressed in either units equals $0.040562n$ where n is the number of non-missing C-peptide concentrations in the meal-test.

Validated macros (ISR and ISR_IML2) have been provided by Novartis and will be used to derive the ISR in █████ production and validation programs.

Calculation of AUC

AUC of ISR/G is defined as $\text{ISR relative to glucose 0-2h (pmol/min/m}^2\text{/mM)} = \text{AUC}_{0-2\text{hr}}$ of $\text{ISR/AUC}_{0-2\text{hr}}$ of glucose. (Glucose is taken from the meal-test laboratory data).

For the patients participating in the meal-test, the laboratory parameters glucose, insulin and C-peptide are also to be measured at the following 7 time-points (minutes in relation to the meal time): -20min, 0min (immediately prior to meal ingestion), 15min, 30min, 60min, 90min, 120min. Let L_1, L_2, \dots, L_7 denote the lab measurements at these 7 time-points (i.e., L_1 is the value measured at 20min pre-meal, L_2 is the value measured at 0min pre-meal, ..., L_7 is the value measured at 2h after the meal) and T_1, T_2, \dots, T_7 be the corresponding times (actual times in units of h), the $\text{AUC}_{0-2\text{hr}}$ is calculated (using the trapezoidal rule) as:

$$\text{AUC}_{0-2\text{h}} =$$

$$[(L_2 + L_3) \times (T_3 - T_2) + (L_3 + L_4) \times (T_4 - T_3) + (L_4 + L_5) \times (T_5 - T_4) + (L_5 + L_6) \times (T_6 - T_5) + (L_6 + L_7) \times (T_7 - T_6)] / 2.$$

Note for above calculation:

- 1) L_1 and T_1 are not used in the calculation. However, if L_2 (the value at 0 min) is missing, the L_1 (the value at -20 min) will be used in its place, but T_2 should still be used in the calculation (i.e., no substitution of T_2 with T_1 is allowed). If the -20 min AND 0 min values are missing, the fasting value (i.e. the standard non meal-test value) will be used to impute the 0 min value. If the 0 minute ISR is missing due to missing 0 minute c-peptide, then the missing c-peptide value should be imputed using this approach, in order to obtain a non-missing 0 minute ISR value.
- 2) If L_i (a lab value) is not missing, but the corresponding time T_i is missing, T_i will be imputed by the planned time (or scheduled time).
- 3) If L_i is missing at a timepoint T_i , the next available value L_{i+1} and timepoint T_{i+1} will be used for calculating AUC.
- 4) AUC_{0-2h} will only be calculated for patients for whom there are at least 3 timepoints (including one of the first two timepoints (i.e., 0 or -20 min pre-meal), the last timepoint (i.e., 2 hours post-meal), and a timepoint in between) where the actual laboratory values are known.

2.2.3.2 Variable definition

The following secondary [REDACTED] efficacy variables will be analyzed. They are grouped by type of analyses for the convenience of describing the related analysis methods.

Variables related to time to event method

- Time to insulin initiation.

Variables related to the rate of change of β -cell function and insulin sensitivity

- Rate of loss of β -cell function (assessed using the AUC of ISR/G) from Visit 4 (week 13) to:
i) end of Period 1; ii) end of Period 2; iii) EOS.
- Rate of change in insulin sensitivity (assessed using the OGIS) from Visit 4 (week 13) to: i) end of Period 1; ii) end of Period 2; iii) EOS.

Variables related to the rate of loss in glycemic control over time

- Rate of loss in glycemic control in HbA1c from 26 weeks after the start of Period 2 to end of Period 2.
- Rate of loss in glycemic control in FPG, i) from Visit 5 (week 26) to: i) end of Period 1; ii) from 26 weeks after start of Period 2 to end of Period 2.

Variables related to change from baseline to endpoint

- Change in AUC of ISR/G (as an assessment of β -cell function) from baseline to: i) end of Period 1; ii) end of Period 2; iii) EOS.
- Change in OGIS (as an assessment of insulin sensitivity) from baseline to: i) end of Period 1; ii) end of Period 2; iii) EOS.

[REDACTED]



2.2.3.3 Statistical analysis methods

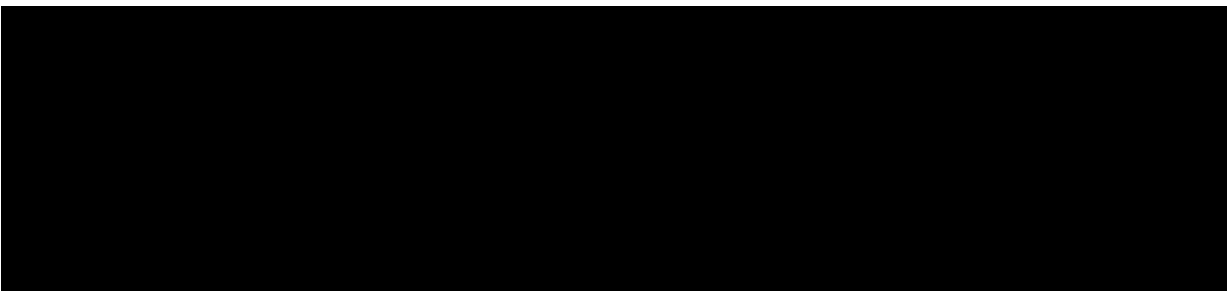
Rate of loss in glycemc control over time

The rate of loss in glycemc control over time will be estimated by the slope of HbA1c over time (in years) as a random coefficient in a linear mixed effect model: the model will be fitted to HbA1c data collected from Week 26 and onwards, up to and including the end of the Period 1 visit, i.e. up to and including the initial treatment failure date. It includes treatment approach, geographic region, baseline HbA1c, time (of HbA1c measurements, in years) and the interaction of treatment approach by time as the fixed effects, and time and intercept as a random effects. Treatment approach and geographic region will be considered as classification variables while baseline HbA1c and time are the covariates. The unstructured covariance will be used as the covariance structure within patients. The actual hypothesis will be tested using the treatment approach by time interaction term. Note that it is not an issue that patients who ‘fail’ at Week 26 will only contribute one data point to the analysis, as this is expected in only a few patients.

The mean slopes within each treatment approach and the difference in mean slopes between two treatment approaches as well as the p-value obtained from the test using the above model will be presented. Graphical representation will also be produced as required.

The analysis will be performed using the FAS. It will be performed in the PPS as supportive analysis as well.

The assumptions for inference using the model as described above (i.e., errors are independently normally distributed with constant variance) will be checked using CSR Section 16 table(s) by subjective examination of the presented residual plots: normal probability plot and residuals versus fitted values plots.





Variables related to the rate of change of β -cell function and insulin sensitivity over time

The rate of change of β -cell function and insulin sensitivity (assessed using the AUC of ISR/G and OGIS respectively) from Visit 4 (week 13) to the end of Period 1, or from Visit 4 (week 13) to the end of Period 2, or from Visit 4 (week 13) to the end of study will be assessed using a similar random coefficient linear mixed effect model used in the primary efficacy endpoint 'rate of loss in glyceic control in HbA1c from Visit 5 (Week 26) to the end of Period 1'.

This analysis is performed in the FAS only. Data collected up to the end of Period 1, or the end of Period 2, or to the end of the study will be used in the analysis for assessing the rate of change from Visit 4 to end of Period 1, or from Visit 4 to end of Period 2, or from Visit 4 to the end of the study respectively.

A general downward trend from Visit 4 to the end of study can be assumed for these parameters.

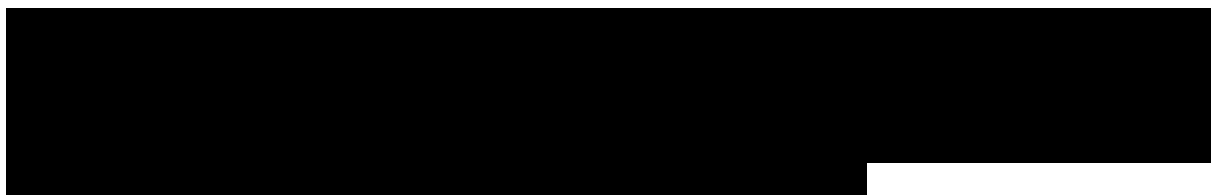
Variables related to the rate of loss in glyceic control over time

The rate of loss in glyceic control in HbA1c or FPG from 26 weeks after the start of Period 2 to the end of Period 2 will be assessed for patients who start insulin therapy in Period 3 (i.e. for treatment intensification) or discontinue the study during Period 2 due to not being able or unwilling to initiate insulin therapy in Period 3. The random coefficient linear mixed effect model used in the primary efficacy endpoint 'rate of loss in glyceic control in HbA1c from Visit 5 (Week 26) to the end of Period 1' will be used in this secondary variable. Only data collected between initiation of Period 2 to insulin initiation on patients who take insulin therapy in Period 3 (i.e. for treatment intensification) or who discontinued the study in Period 2 due to not being able or unwilling to initiate insulin therapy in Period 3 will be included in this model. Patients who complete the study in Period 1 or Period 2 will not be included in the analysis. The analysis will be performed in the FAS only. Note that it is not known how many patients will proceed into Period 3, so may need to later review whether the planned analysis for the subgroup of patients entering Period 3 (or being unable/unwilling to) is sensible.

The rate of loss in glyceic control in FPG from 26 weeks to the end of Period 1 will be assessed using the same random coefficient linear mixed effect model used in the primary efficacy endpoint 'rate of loss in glyceic control in HbA1c from Visit 5 (Week 26) to the end of Period 1'. This analysis is performed in the FAS only.

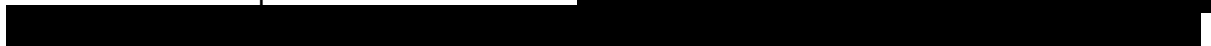
Variables related to change from baseline to endpoint

An analysis of covariance (ANCOVA) model with treatment approach, geographic region as classification variables and baseline value as a covariate will be used to assess all secondary efficacy variables related to change from baseline to endpoint. The least squares mean (“adjusted mean”) change from baseline for each treatment approach, the difference in the least squares mean changes between the two treatment approaches, and the two-sided adjusted 95% confidence interval along with the p-value for the difference will be obtained from this analysis model and presented.



Supportive descriptive analysis for secondary [REDACTED] variables over time

Summaries of absolute values and change from baseline to: i) end of Period 1; ii) end of Period 2; iii) EOS; in secondary [REDACTED] (including FPG, [REDACTED], AUC of ISR/G, OGIS, [REDACTED]) by treatment approach and visit will be presented in the FAS.



2.3 Safety evaluation

The safety and tolerability of vildagliptin 50mg bid as add-on to metformin over 5-years of treatment will be compared to placebo as add-on to metformin based on data collected during the treatment period in SAF. All analyses will use all available data collected during the entire study. In addition, key safety analyses will be performed on data obtained during Period 1 only. Patients will be analyzed according to the treatment approach they received (vildagliptin 50mg bid + metformin, placebo + metformin).

The incidence of treatment emergent adverse events (events started after the first dose of randomized study medication or events present prior to start of randomized treatment but increased in severity) will be summarized by primary system organ class, preferred term, severity and relationship to study drug. Any event captured on the database will be included in AE summary tables/listings.

In order to assess safety and tolerability of vildagliptin as compared to placebo, key safety variables (overall AEs, SAEs, AEs leading to study drug discontinuation or study drug interruption, incidence of hypoglycemia, predefined AE risks, predefined categories of liver enzyme (ALT/AST) and CPK and persistent elevations) will also be summarized by treatment approach for the SAF.

The number and percentage of run-in failures reporting any adverse events, serious adverse events or adverse events leading to study drug discontinuation will also be summarized. Note

these events will not be included in other AE outputs, as these patients will not be included in SAF.

2.3.1 Overall experience of adverse events (AEs)

The number and percentage of patients reporting any adverse event during the treatment period of the study will be summarized by primary system organ class, preferred term and treatment approach overall (including study periods 1, 2 and 3) and by study period. Percentages will be based on the number of patients in the SAF per treatment approach and study period. An AE will be assigned to a period if its start date is after or on the start date but not later than the end date of that period. All AEs will be assigned to a specific study period – if it is not possible to assign an AE into a specific study period of the patient based on the AE start date, then the event will be assigned to the last study period started by the patient. This will ensure that all AEs will be categorized to a specific study period

The most common adverse events reported ($\geq 2\%$ in any group) during the treatment period of the study will be presented in descending frequency starting from the most common event in a Text Table. A complete table of all events reported in descending frequency will be prepared..

Adverse events that were suspected to be related to study drug will be summarized. Note that if the relationship to study drug is missing for any adverse event, then that adverse event will be assumed to be related to study drug.

The maximum severity of all adverse events during the treatment period of the study will be summarized. Note that if the severity is missing for any adverse event, then that adverse event will be assumed to be severe.

An overall summary of the number and percentage of patients with drug-related AEs, severe AEs, AEs with outcome of death, SAEs and adverse events leading to permanent discontinuation of treatment will be presented by study period and treatment approach.

The number and percentage of patients reporting any adverse event during the treatment period of the study will be summarized by standardized MedDRA query (SMQ) level and treatment approach as well.

In addition, the number and percentage of patients reporting any adverse event during the treatment period will also be summarized by primary system organ class, preferred term, treatment approach and age group (as defined in section 2.1.5)) overall and by treatment period..

Note that there will be no AE adjudication for this study.

2.3.1.1 Deaths, other serious and other significant adverse events

An overall summary of the number and percentage of deaths, serious adverse events (SAEs), and adverse events (including laboratory abnormalities) leading to discontinuation, study drug interruption or considered to be clinically significant will be presented by study period and treatment approach.

Deaths

The number and percentage of deaths will be summarized by study period and treatment approach and the principal cause in primary system organ class and preferred term.

Serious adverse events (SAEs)

The number and percentage of patients with serious adverse events (SAEs) will be summarized by primary system organ class, preferred term, study period and treatment approach.

Other significant adverse events – adverse events leading to discontinuations

The number and percentage of patients with adverse events that lead to discontinuation will be summarized by primary system organ class, preferred term, study period, and treatment approach.

Other significant adverse events – adverse events leading to dose adjustments or study drug interruptions

The number and percentage of patients with adverse events that lead to dose adjustments or study drug interruptions will be summarized by primary system organ class, preferred term, study period, and treatment approach.

Other significant adverse events – predefined adverse event risks

The number and percentage of patients with adverse events that were predefined risks (see [Section 1.6](#) for definitions) will be summarized by preferred term, predefined risk category, maximum severity, study period, and treatment approach.

2.3.2 Laboratory values

The details of laboratory parameters collected, SI units (note, only SI units will be reported), and normal ranges will be listed in [Section 16 of CSR](#). Note that lipid parameters are summarized along with other biochemistry laboratory parameters. Meal test laboratory parameters (glucose, insulin and c-peptide) are for use in efficacy analyses only and so should not be included in safety laboratory outputs.

Notable ranges (absolute) criteria of selected laboratory variables are defined in [Section 1.8](#) above. The laboratory values to be considered for notable ranges or relevant changes will include all post-baseline values up to the final scheduled visit, including values at unscheduled visits prior to the last visit. Only treatment emergent notable abnormalities will be considered (i.e. notably abnormal measurements which were not present at any pre-baseline visit). In addition to summary tables, a listing of patients with any treatment emergent notable laboratory results during the treatment period will be provided by treatment approach and patient number, along with their values at all visits, including repeated, unscheduled and follow-up laboratory values. In the listing, all abnormal values outside of normal range and notable values will be flagged, e.g., L/H for a value being below/above normal range, LN/HN for a value being in notable low/high range.

For each numerical laboratory parameter, the laboratory values will be reported in listings to the same precision as that for the unit in UNITCONV dataset. The precision to use in tables is as per [Section 1.3.1](#).

For continuous variables with a value <lower limit, these will be imputed as being half of the lower limit. For continuous variables with a value >upper limit, these will be imputed as being equal to the upper limit.

2.3.2.1 Hematology

The values of hematology laboratory test results at baseline and endpoints (End of Period 1, End of Period 2, End of Study) will be summarized, as well as the changes from baseline to each endpoint.

The maximum/minimum value obtained during the treatment period, and the change from baseline to maximum/minimum value will be summarized by treatment approach for each hematology parameter.

The number and percentage of patients with abnormal hematology test results meeting notable criteria post-baseline, which were not present prior to randomization, will be tabulated. This summary will be repeated for abnormalities occurring in Period 1.

2.3.2.2 Biochemistry

The values of biochemistry laboratory test results at baseline and endpoints (End of Period 1, End of Period 2, End of Study) will be summarized, as well as the changes from baseline to each endpoint.

The maximum/minimum value obtained during the treatment period, and the change from baseline to maximum/minimum value will be summarized by treatment approach for each hematology parameter.

The number and percentage of patients with abnormal biochemistry test results meeting notable criteria post-baseline, which were not present prior to randomization, will be tabulated. This summary will be repeated for abnormalities occurring in Period 1.

The number and percentage of patients with hepatic enzyme elevations (defined in [Table 1-3](#)), and persistent hepatic enzyme elevations will be presented, and the patients meeting any of the criteria will be listed. Note that for this study total bilirubin will be used for all criteria involving bilirubin. If at any visit total bilirubin has not been measured then the elevation will not be assessed for that visit (even if direct bilirubin has been measured).

The liver panel A testing data for patients with hepatic enzyme elevations will be listed at patient level. Liver panel B testing data will not be considered as part of the RAP, but will be included in narratives if applicable.

2.3.2.3 Urinalysis

The baseline and endpoints (End of Period 1, End of Period 2, End of Study) values of categorical urinalysis test results will be summarized in number and percentage by treatment approach. The baseline and endpoints (End of Period 1, End of Period 2, End of Study)

values of continuous urinalysis test results will be summarized by treatment approach using descriptive statistics.

2.3.3 Vital signs

Baseline and endpoint values for blood pressure (sitting), and pulse (sitting) will be presented. Summary statistics for vital signs data will be presented by study visit. The summary by study visit will include only values obtained at the scheduled visit.

The number of patients with notably abnormal vital signs [REDACTED] post-baseline will be presented. This summary will be repeated for abnormalities occurring in Period 1. The criteria for notable abnormalities (see [Section 1.9](#)) will be listed.

2.3.4 Electrocardiograms (ECGs)

Overall interpretation and details of abnormalities are collected at Screening and selected visits.

Change from the baseline ECG result to the endpoint (End of Period 1, End of Period 2, End of Study) ECG result will be summarized by a shift table. The number of patients with clinically significant abnormalities or not at baseline and at each endpoint ECG assessment will be tabulated.

2.3.5 Hypoglycemic events

Hypoglycemic events will be characterized by event profile, such as ability to self-treat, self monitoring of plasma glucose level, precipitating event, time from last meal, time from last dose, and time of the day.

The incidence of hypoglycemic events (coded as hypoglycemia) and asymptomatic low blood glucose (coded as Blood Glucose Decreased), entered into the Glycemia study diary eCRF and meeting the event criteria defined in [Table 1-1 \(Section 1.7\)](#), are included in all of the AE summaries in [Section 2.3.1](#).

In addition, patients reporting at least one hypoglycemic event, the subgroup of reporting 1 event, 2 events, and >2 events, patients discontinued due to hypoglycemic events, patients reporting grade 2 events, and patients reporting suspected grade 2 events will be summarized by numbers and percentages for each treatment approach overall (including study periods 1, 2 and 3) and by study period. The summary will be repeated by age group (as defined in [section 2.1.5](#)).

The hypoglycemic events will further be summarized by event profile as follows:

- plasma glucose level ≤ 2.2 , $>2.2-\leq 2.8$, $>2.8-\leq 3.1$ and ≥ 3.1 mmol/L, or Not recorded; note that if a patient is unable to initiate self-treatment and plasma glucose is not recorded, this is graded as a suspected grade 2 event)
- grade (grade 1, grade 2 or suspected grade 2)
- symptom (in preferred term)
- precipitating event (None, Missed/delayed meal, Strenuous exercise, Alcohol consumption, Other or Not recorded)

- severity (Mild, Moderate, Severe)
- relationship to study drug (Not suspected, Suspected)
- action taken (Study drug dosage adjusted / temporarily discontinued, Study drug permanently discontinued, Concomitant medication taken, Non-drug therapy given (includes oral carbohydrates), Hospitalization / prolonged hospitalization, no action taken)
- time interval between last meal and event (0–4, >4–6, >6 hours, not recorded),
- time interval between last dose and event (0–4, >4–6, >6 hours, not recorded),
- time of the day in 24-hour clock (>00:00–06:00, >06:00–12:00, >12:00–18:00, or >18:00–24:00),

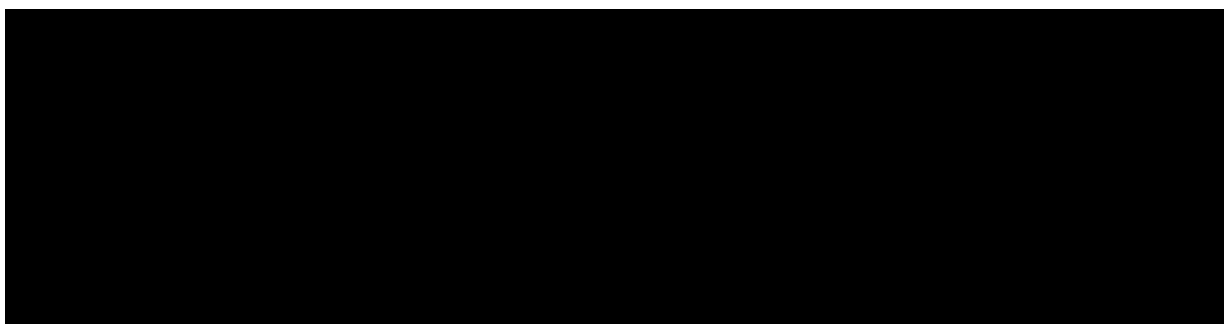
Asymptomatic blood glucose events will be similarly summarized by event profile and the number of patients with at least one asymptomatic low blood glucose measurement will also be summarized.

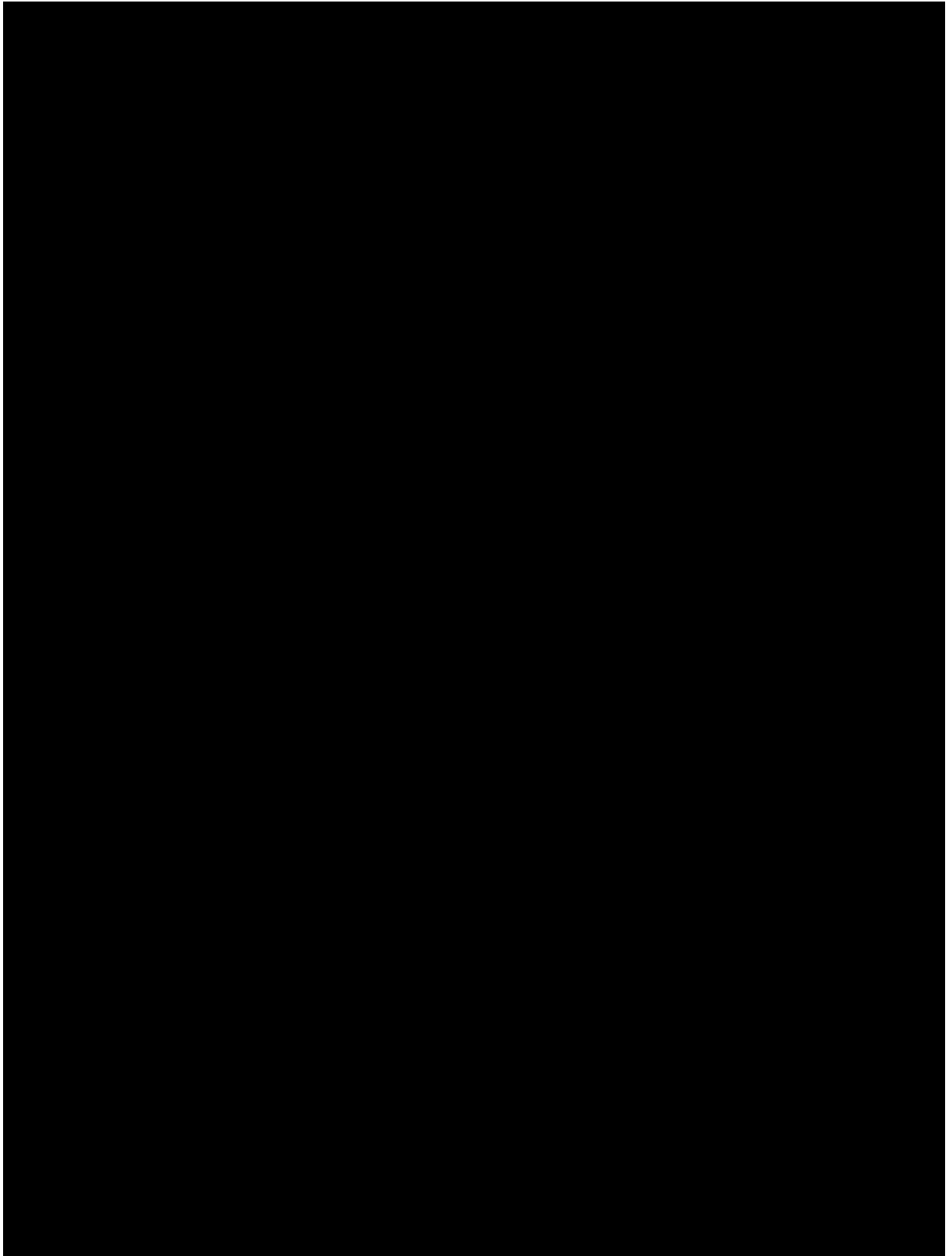
A listing of hypoglycemic events and asymptomatic low blood glucose events will be presented by treatment approach and patient number.

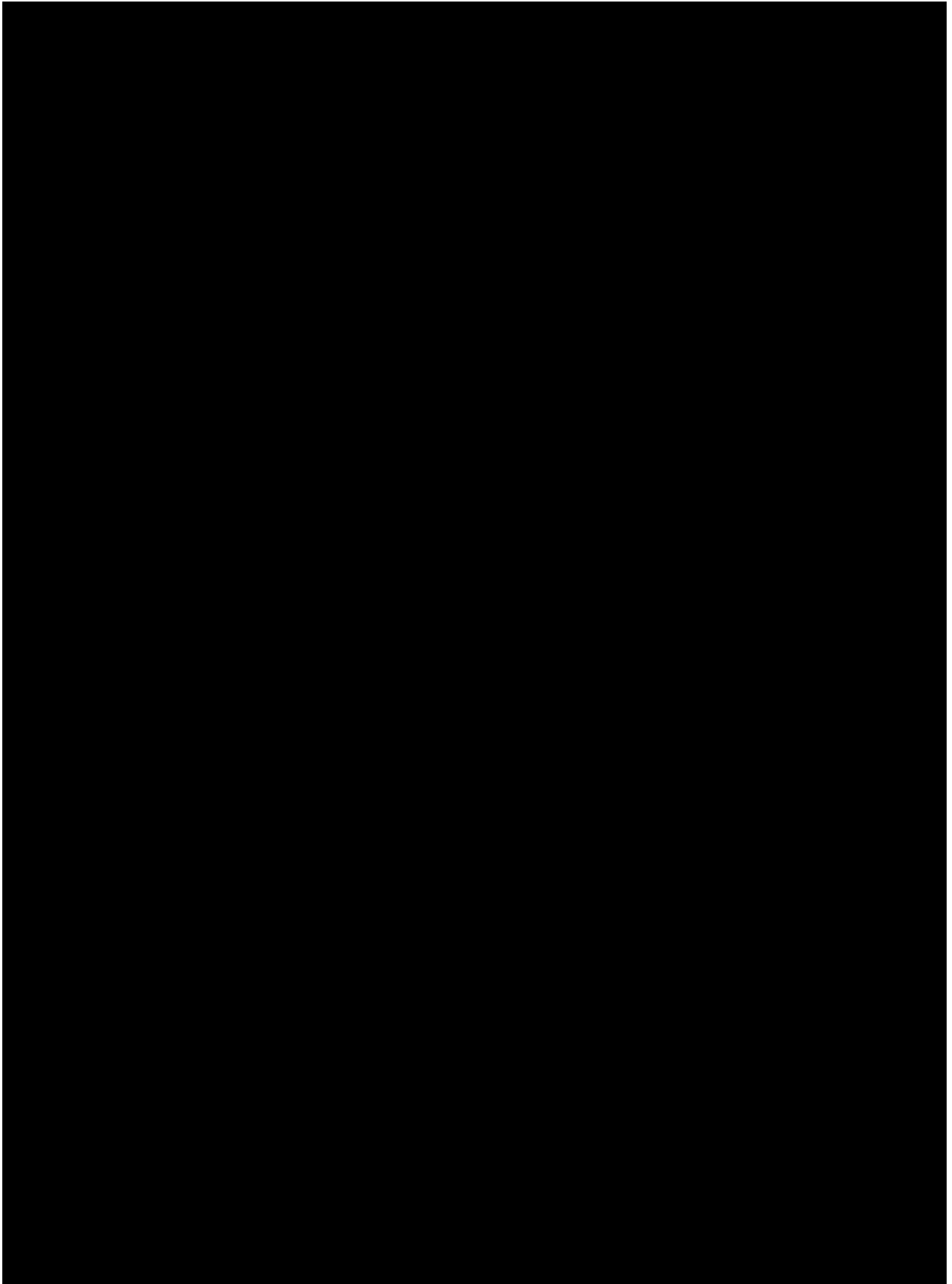
Note: As defined in [Table 1-1](#), for patients who are able to initiate self-treatment (where necessary) symptoms suggestive of hypoglycemia which are not confirmed by glucose measurements will be recorded in the Glycemia diary eCRF. Since these are not confirmed hypoglycemia events, these symptoms will not be reported in the hypoglycemia summary tables and listings, these symptoms will appear only in the AE tables and listings. For patients unable to initiate self-treatment (where necessary) symptoms suggestive of hypoglycemia that were not confirmed by glucose measurements will be classed as suspected hypoglycemic events and will therefore be reported in the hypoglycemia summary tables and listings, as well as the AE listings.

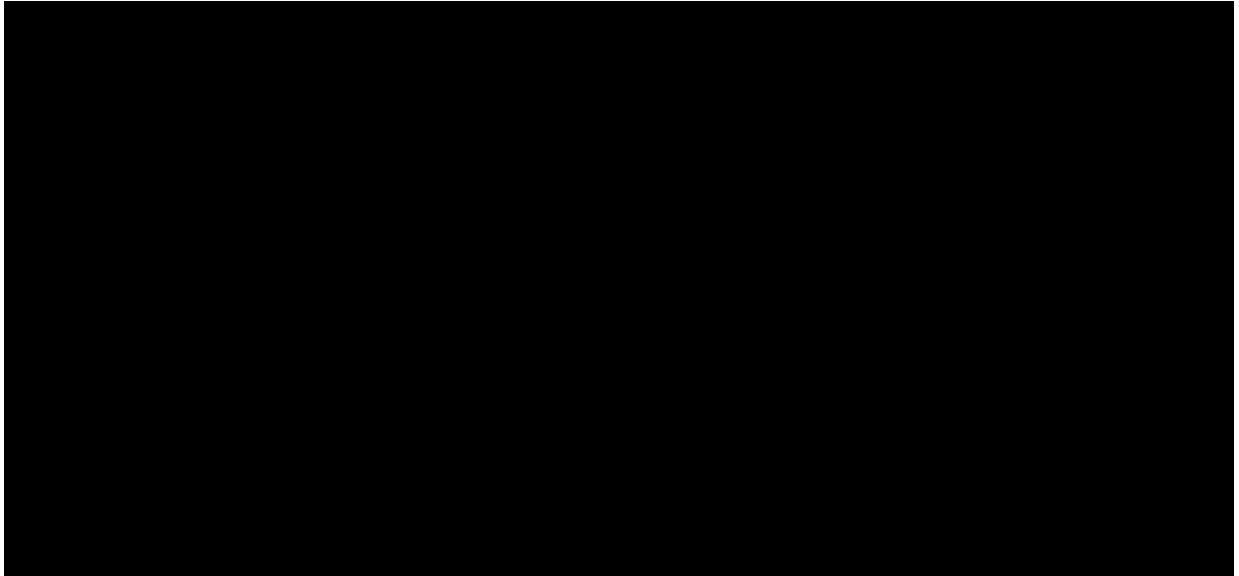
Note: Self-monitoring glucose can be recorded in the database as either plasma or whole blood. However, only plasma glucose will be reported in the tables and listings. Thus blood glucose values will be converted into plasma glucose equivalent, which can be done by multiplying a conversion factor of 1.12, i.e., plasma glucose = blood glucose × 1.12.

The incidence of hypoglycemia and severe hypoglycemia during Period 1 will be compared for the SAF using a chi-square test between the two treatment approaches for all patients provided that the incidence such these events is more than 5 in each treatment approach.









2.5 Interim analyses

No interim analyses are planned for this study.

2.6 Other topics

No other topics will be studied.

2.7 Determination of sample size

A total sample size of 1000 randomized patients per treatment arm (in 1:1 allocation ratio to vildagliptin + metformin and metformin monotherapy) is planned.

The sample size calculation assumed that all randomized patients are to be followed up for 5 years unless patients dropped out from the study for various reasons (lack of efficacy, AEs, abnormal labs, lost to follow-up etc.), and that the yearly dropout rate is 11% (estimate based on ADOPT data (Kahn S et al, 2006)).

The existing vildagliptin study data suggest that approximately 10% of vildagliptin patients will have an HbA1c >7.0% after the first 3 months of the study (initial response phase), since those patients who are randomized with an HbA1c measurement above the failure threshold (7.0%) may never have an HbA1c measurement below the required threshold during the study and will therefore be counted as failures during the first 13 weeks. A similar proportion is assumed for the comparator arm. Hence it is expected that the difference in failure rate is likely to be small early in the study, but will diverge as the study progresses. The power calculations have been adjusted to take this assumption into account using statistical simulations.

The simulations showed that assuming an annual initial treatment failure rate of 7.1% in the metformin monotherapy arm (estimated based on ADOPT data), incorporating a 10% initial failure rate after 13 weeks in each treatment arm (due to some patients with baseline HbA1c $\geq 7.0\%$), 1000 patients per treatment arm would be sufficient to detect a hazard-ratio of 0.75 between vildagliptin + metformin and metformin alone (corresponding to a risk reduction rate of 25% in vildagliptin + metformin group versus metformin alone) with approximate 75% power and a 1-sided significance level of 0.025 (corresponding to a 2-sided test at 0.05).

2.8 Changes in the conduct of the study or planned analyses


The following change in the analysis is proposed in protocol amendment 3 (4th Oct 2016) which defers with the analysis proposed in this document. These proposed changes have come about through development of the RAP modules and discussions around the analyses.

- Treatment start date to be defined as the earliest vildagliptin/placebo date, as recorded on the DAR page of the eCRF (i.e. do not use visit 3 date). Consistently use this definition for day 1 and durations related to day 1.
- '+1' should be included in duration of exposure calculation when last study drug date is available.
- '+1' should not be included in duration of exposure calculation when last study drug date is missing/incomplete.
- Further detail added to better define the analysis sets.
- Geographic region to be included as a classification variable in the analysis of time to initial treatment failure (and all similar analyses).
- In the analysis of rate of loss in glycaemic control over time, time will be included as a fixed effect and as a random effect.
- In the analysis of time to insulin initiation, patients who complete the study in Period 1 or Period 2 will be censored at their last study visit.
- OGIS formula revised as per 23135 study.
- Due to the large number of patients (patients will be recruited until we have the required amount) and the robustness of analyses methods we can assume we will have sufficient patients data for the beta-cell function, insulin sensitivity, AUC of ISR/G and OGIS analyses.

█ [REDACTED]

█ [REDACTED]

- The definition of initial treatment failure is revised to be the date of the second of the two consecutive measurement of HbA1c value $\geq 7\%$. The rationale of this change is that this will be consistent with the Period 1 end date definition. The decision of this definition update is documented in the RAP II meeting minutes in RAP Module 1.
- Main analysis approach includes only one primary endpoint in this SAP. As the probability and risk of time to initial loss of glycaemic control as primary efficacy variable is presented as a hazard ratio and failure rate over time, the statistical approach and power calculations (including the impact of the known retention) was updated accordingly while moving the assessment of the rate of loss of glycaemic control by an annualised slope of mean HbA1c over time from week 26 to the end of period 1 to be a key secondary analysis instead.

- Baseline characteristics driving the pre-planned sub-analyses to a previously unpredictable direction, is clearly documented in the SAP to ensure transparency and external validity of the study
- 
- Time to second treatment failure and the subgroup analysis of the primary endpoint is added as a post-hoc analysis which is to be included in the CSR.

References

[ADA-EASD Consensus Statement \(2009\)](#) Medical management of hyperglycaemia in type 2 diabetes mellitus: a consensus algorithm for the initiation and adjustment of therapy. *Diabetologia*; 52:17-30.

[Hochberg and Tamhane \(1987\)](#) Multiple Comparison Procedures. Wiley Series.

[Mari A, Ludvik B, Pacini G, et al \(2001\)](#) A model-based method for assessing insulin sensitivity from the oral glucose tolerance test. *Diabetes Care*; 24: 539-548.

[Rigalleau V, Lasseur C, Perlemoine C, et al \(2005\)](#) Estimation of glomerular filtration rate in diabetic subjects. Cockcroft formula or modification of diet in renal disease study equation? *Diabetes Care* 28:838-843.

[REDACTED]

[Scherbaum W, et al. \(2008\)](#) Evidence that vildagliptin attenuates deterioration of glycaemic control during 2-year treatment of patients with type 2 diabetes and mild hyperglycaemia. *Diabetes Obes Metab*; 10:1114–1124.

[REDACTED]

[SAS/STAT User's Guide](#), version 9.2, SAS Institute, Inc., Cary, North Carolina.