Cover Page for Statistical analysis plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT03268005
Sponsor trial ID:	NN1218-4113
Official title of study:	Efficacy and Safety of Fast-acting Insulin Aspart Compared to NovoRapid® both in Combination with Insulin Degludec with or without Metformin in Adults with Type 2 Diabetes (onset® 9)
Document date:	20 June 2019

16.1.9 Documentation of statistical methods

List of contents

Redacted statistical analysis plan
Includes redaction of personal identifiable information only.

.

Statistical Analysis Plan

Trial ID: NN1218-4113

Efficacy and Safety of Fast-acting Insulin Aspart Compared to NovoRapid® both in Combination with Insulin Degludec with or without Metformin in Adults with Type 2 Diabetes (onset® 9)

Author:

[redacted]

Table of contents

Page

List of abbreviations

ADA	American Diabetes Association
AE	adverse event
ALT	alanine aminotransferase
AST	aspartate aminotransferase
BG	blood glucose
CI	confidence interval
ECG	electrocardiogram
FAS	full analysis set
FPG	fasting plasma glucose
HbA_{1c}	glycosylated haemoglobin
HDL	high density lipoprotein
ICH	international council on harmonisation
LDL	low density lipoprotein
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
PG	plasma glucose
PP	per protocol
PPG	postprandial glucose
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SMPG	self-measured plasma glucose
T2DM	type 2 diabetes mellitus
TEAE	treatment emergent adverse event

1 Introduction

1.1 Trial information

This is a phase 3b, 16-week, multicentre, multinational, 1:1 randomised, double-blind, active controlled, treat-to-target, parallel group trial with a 12-week run-in period comparing the effect and safety of faster aspart to NovoRapid® both in combination with insulin degludec with or without metformin in subjects with T2DM treated with a basal-bolus regimen.

The total duration of the trial is approximately 34 weeks divided into the following periods (see Figure 1–1):

- An approximately 2-week screening period
- A 12-week run-in period primarily for optimisation of the basal insulin and subject training
- A 16-week treatment period
- A 30-day follow-up period: FU1; 7 days after end of treatment and FU2; 30 days after end of treatment

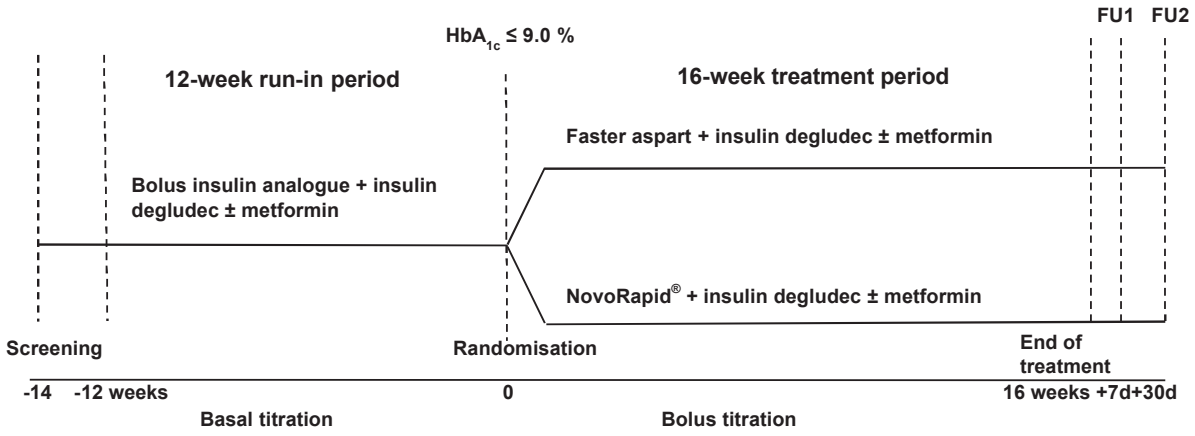


Figure 1–1 Trial design

The trial includes a screening period followed by weekly visits/phone contacts during the trial. At Visit 2, all eligible subjects will be enrolled in a 12-week run-in period. After the run-in period, subjects eligible for randomisation ($HbA_{1c} \leq 9.0\%$ measured at Visit 13) will be randomised (1:1) to receive double blinded treatment with either faster aspart or NovoRapid® both in combination with once daily insulin degludec with or without metformin.

Subjects will have a standardised meal test at baseline (Visit 14 before randomisation) and at end of treatment (Visit 30). The meal test is described in more detail in section 8.3.1 in the protocol.

After the 16-week treatment period, each subject will have a 30-day safety follow-up period.

The primary objective is to confirm the effect in terms of glycaemic control of treatment with faster aspart compared to NovoRapid® both in combination with insulin degludec with or without metformin in adults with T2DM treated with a basal-bolus regimen, using a non-inferiority approach.

The secondary objectives are to confirm superiority of faster aspart compared to NovoRapid® both in combination with insulin degludec with or without metformin in adults with T2DM treated with a basal-bolus regimen in terms of:

- Postprandial glucose regulation
- Overall glycaemic control
- Postprandial glucose excursions

To compare the safety of faster aspart to NovoRapid® both in combination with insulin degludec with or without metformin in adults with T2DM treated with a basal-bolus regimen.

Subjects prematurely discontinued from trial product should continue with the per protocol planned visits at 4 (Visit 18), 8 (Visit 22), 12 (Visit 26), 16 (Visit 30) weeks after randomisation depending on when the subject discontinues trial products.

For further details on handling of subjects that prematurely discontinue from trial product and the trial in general, please see the trial protocol.

1.2 Scope of the statistical analysis plan

This statistical analysis plan (SAP) is based on the protocol *Efficacy and Safety of Fast-acting Insulin Aspart Compared to NovoRapid® both in Combination with Insulin Degludec with or without Metformin in Adults with Type 2 Diabetes (onset® 9)*, version 4.0 (dated 23 February 2018). The statistical analyses and derivations of endpoints presented in this SAP are almost identical to those described in the protocol. It contains minor clarifications for derivations, calculation of endpoints and analyses as well as some additions.

The changes to the statistical considerations proposed in this SAP and the reasons for the changes are described in section 3 and will be reported in the clinical trial report (CTR).

2 Statistical considerations

General considerations

In general, for endpoints evaluated as a change from baseline and/or where a baseline adjustment is made, baseline is defined as information collected at randomisation (Visit 14). In case a measurement is not available at randomisation, the most recent measurement prior to randomisation will be used as baseline.

Two observation periods are defined, "in-trial" and "on-treatment", and it will be specified which period each analysis will use.

- In-trial: The observation period from date of randomisation and until last trial -related subject-site contact. The in-trial observation period includes data collected after treatment discontinuation.
- On-treatment: The observation period from date of first dose of randomised NovoRapid®/faster aspart and to 7 days after the first occurrence of:
    - The day of last dose of randomised NovoRapid®/faster aspart
    - The day before initiation of ancillary treatment

All efficacy endpoints will be summarised and analysed using the full analysis set (FAS), unless otherwise stated. Safety endpoints will be summarised using the safety analysis set and analysed using the FAS, unless otherwise stated. The FAS and safety analysis set are defined in section 2.2.

Presentation of results from a statistical analysis will include the estimated mean treatment effects (LSMeans) for change from baseline, if applicable. Estimated mean treatment differences (or ratios) will be presented together with two-sided 95% confidence interval (CI) for all endpoints analysed statistically.

Data collected before randomisation (Visit 14) will only be summarised descriptively.

Testing strategy and estimands

The primary objective of the trial is to confirm the effect of treatment with faster aspart compared to NovoRapid®, both in combination with insulin degludec with or without metformin in adults with T2DM in terms of glycaemic control, using a non-inferiority approach.

More specifically the upper limit of the 95% confidence interval for the difference between faster aspart and NovoRapid® should be compared to a non-inferiority margin of 0.4%. If it is below or equal to 0.4% non-inferiority will be considered established and effect demonstrated.

The trial also aims to confirm superiority of treatment with faster aspart for a number of secondary confirmatory endpoints. The family-wise type I error rate will be controlled in the strong sense

using a hierarchical (fixed sequence) testing procedure under the framework of the primary estimand. This is based on a priority ordering of the null-hypotheses, and testing them in this order using the two-sided 95% confidence interval approach until an insignificant result appears. The effect is that rejection of a null hypothesis only will be considered for analyses where all previous null-hypotheses have been rejected in favour of faster aspart.

The steps in the hierarchical testing procedure are as follows:

Step 1 (Primary analysis): HbA_{1c} non-inferiority of faster aspart versus NovoRapid®

Step 2: 1-hour PPG increment (meal test) superiority of faster aspart versus NovoRapid®

Step 3: HbA_{1c} superiority of faster aspart versus NovoRapid®

Step 4: 1,5-anhydroglucitol superiority of faster aspart versus NovoRapid®

Primary estimand

The primary estimand is defined as the treatment difference between subjects randomised to faster aspart and NovoRapid® both in combination with insulin degludec with or without metformin, in adults with T2DM not optimally controlled with a basal-bolus regimen assessed by change from baseline in HbA_{1c} 16 weeks after randomisation for all randomised subjects regardless of treatment discontinuation or use of ancillary treatment.

The primary estimand assesses the expected benefit a future population with T2DM can achieve if prescribed to faster aspart as compared to NovoRapid®. By not putting any restrictions on the treatment adherence, this estimand aims at a difference as close as possible to the one that can be expected in real-world clinical practice, provided that the treatment adherence and use of ancillary treatment reflects clinical practice. Thereby the primary estimand provides a treatment difference for clinicians concerning the glycaemic effect of faster aspart compared to NovoRapid® in the day to day life in subjects with T2DM in an adult population.

Secondary estimand

A secondary estimand is defined as the treatment difference in change from baseline in HbA_{1c} 16 weeks after randomisation between treatment with faster aspart and treatment with NovoRapid® both in combination with insulin degludec with or without metformin in adult subjects with T2DM not optimally controlled with a basal-bolus regimen if all subjects had adhered to randomised treatment and did not receive ancillary treatment.

The condition ‘adhered to randomised treatment and did not receive ancillary treatment’ should be interpreted as the exclusion of information collected after initiation of antidiabetic treatment that can mask or exaggerate the effect of the initially randomised treatment. Only data collected prior to

discontinuation of trial product or initiation of ancillary treatment is used to draw inference. This avoids confounding effects of ancillary treatment.

The two estimands will be repeated for the confirmatory endpoints

- Change from baseline in 1-hour PPG increment (meal test) 16 weeks after randomisation
- Change from baseline in 1,5-anhydroglucitol 16 weeks after randomisation

Visit reallocation

Subjects that prematurely discontinue from treatment or withdraw from trial will attend end of treatment visit called visit 30A. Data collected at this visit will be reallocated to the next scheduled visit where the given assessment is planned. As a general rule, all observed values from randomised subjects will be used in all statistical analyses, but in case two different values are associated to the same visit in time, the use of a given value will depend on the estimand of interest. For the primary estimand the reallocated on-treatment value will not be used and for the secondary estimand the reallocated on-treatment value will be used.

2.1 Sample size calculation

The sample size is determined to ensure a sufficient power for step 1 and step 2 in the hierarchical testing procedure for the primary estimand presented in section 2. The power for step 3 and 4 in the hierarchical testing procedure will also be presented. The sample size is determined using a non-inferiority limit of 0.4% in step 1, which was chosen as described in section 5.2.1 in the protocol. The statistical evaluation will be done as described in sections 2.3 and 2.4.

In previous confirmatory trials where faster aspart has been investigated, the completion rates have been high. Therefore it will be expected that treatment discontinuation might be as low as 10% where trial discontinuation constitutes half of these and with similar withdrawal reasons in the treatment arms.

Power for the non-inferiority step (Step 1) is based on a t-statistic under the assumption of a one-sided test of size 2.5% for the FAS. A mean treatment difference of -0.1% for the comparison between faster aspart and NovoRapid® in favour of faster aspart is expected. As trials in this population where data from treatment withdrawn subjects is retrieved is limited, a conservative estimate of the standard deviation (SD) in change from baseline in HbA_{1c} of 0.8% was chosen. The power for superiority in step 3 will be calculated using the same assumptions as for step 1 but without the non-inferiority margin.

For determination of power in step 2 in the hierarchical testing, where change from baseline in 1-hour PPG increment 16 weeks after randomisation is compared between faster aspart and NovoRapid® a t-statistic with a two-sided test of size 5% is used, where the treatment difference is

expected to be at least 0.6 mmol/L [11 mg/dL]. The SD=3.5 mmol/L [63 mg/dL] in change from baseline in 1-hour PPG increments 16 weeks after randomisation based on laboratory analysed PG in a standardised meal test will be considered reasonable based on trials NN1218-3852 and NN1218-3853.

The power in step 4, where superiority of faster aspart over NovoRapid® in change from baseline in 1,5-anhydroglucitol 16 weeks after randomisation is tested a t-statistic with a two-sided test of size 5% is used. The mean treatment difference is expected to be at least 0.2 µg/mL and an SD of 3.5µg/mL will be used based on trials NN1218-3852 and NN1218-3853.

The power calculations are done using proc power in SAS 9.4. Please refer to Table 2–1 for assumptions for the sample size calculation.

Table 2–1 Specifications assumed for sample size calculation

	Statistical test	Significance level	Analysis population	Non-inferiority margin	SD	Mean difference	Randomisation scheme
Step 1	2-group t-test	One-sided 2.5 %	FAS	0.4 % (absolute)	0.8 %	-0.1 %	1:1
Step 2	2-group t-test	Two-sided 5.0%	FAS	NA	3.5 mmol/L	0.6 mmol/L	1:1
Step 3	2-group t-test	Two-sided 5.0%	FAS	NA	0.8 %	-0.1 %	1:1
Step 4	2-group t-test	Two-sided 5.0%	FAS	NA	3.5 µg/mL	0.2 µg/mL	1:1

In Table 2–2 the sensitivity of the sample size to the power shown for three different sizes of FAS. Three different choices of the mean difference are used to calculate the power in step 2.

Table 2–2 Sensitivity of sample size to power

N total	N per arm	Step 1			Step 2			Step 3			Step 4		
FAS	FAS	Mean diff (%)	SD (%)	Power (%)	Mean diff (mmol/L)	SD (mmol/L)	Power (%)	Mean diff (%)	SD (%)	Power (%)	Mean diff (µg/mL)	SD (µg/mL)	Power (%)
		-0.1	0.8	>99.9	0.5	3.5	58.1	-0.1	0.8	47.4	0.2	3.5	13.9
920	460	-0.1	0.8	>99.9	0.6	3.5	73.8	-0.1	0.8	47.4	0.2	3.5	13.9
		-0.1	0.8	>99.9	0.7	3.5	85.8	-0.1	0.8	47.4	0.2	3.5	13.9
		-0.1	0.8	>99.9	0.5	3.5	64.7	-0.1	0.8	53.4	0.2	3.5	15.5
1072	**536**	**-0.1**	**0.8**	**>99.9**	**0.6**	**3.5**	**80.1**	**-0.1**	**0.8**	**53.4**	**0.2**	**3.5**	**15.5**
		-0.1	0.8	>99.9	0.7	3.5	90.5	-0.1	0.8	53.4	0.2	3.5	15.5
		-0.1	0.8	>99.9	0.5	3.5	70.4	-0.1	0.8	58.9	0.2	3.5	17.0
1224	612	-0.1	0.8	>99.9	0.6	3.5	85.0	-0.1	0.8	58.9	0.2	3.5	17.0
		-0.1	0.8	>99.9	0.7	3.5	93.8	-0.1	0.8	58.9	0.2	3.5	17.0

In conclusion, 1072 subjects in the FAS (536 subjects per group) will ensure a marginal power of >99.9% to show non-inferiority in step 1, given that the actual treatment difference is -0.1%, and a marginal power of 80.1% to show superiority in step 2, given that the actual treatment difference is 0.6 mmol/L.

Assuming a screening failure rate of 30% and run-in failure rate of 15%, 1803 subjects should be screened for inclusion in the trial.

2.2 Definition of analysis sets

The following analysis sets are defined in accordance with the ICH-E9 guidance[1].

- Full Analysis Set (FAS) includes all randomised subjects. In exceptional cases, randomised subjects may be excluded from the FAS. In such cases, the reason for exclusion will be justified and documented. Subjects in the FAS will contribute to the evaluation "as randomised"
- Per Protocol (PP) Analysis Set includes all subjects in the FAS, excluding subjects who:
    - Have violated any inclusion criteria
    - Have fulfilled any exclusion criteria

  Subjects in the PP analysis set will contribute to the evaluation "as treated"
- Safety Analysis Set includes all subjects receiving at least one dose of randomised treatment. Subjects in the safety analysis set will contribute to the evaluation "as treated".

Randomised subjects who are lost to follow-up, and where no exposure information of the trial product or its comparator is available after randomisation, will be handled as unexposed.

Before data are released for statistical review, a blinded review of all data will take place to identify serious non-adherence to the protocol that may potentially affect the results. Furthermore, extreme values and outliers will be identified by the statistician during programming and data review, according to ICH-E9[1]. This will be performed using a fake randomisation.

The subjects or observations to be excluded, and the reasons for their exclusion must be documented and signed by those responsible before database lock. The subjects and observations excluded from analysis sets, and the reason for this, will be described in the clinical trial report.

2.3 Primary endpoint

The primary endpoint is the change from baseline in HbA_{1c} 16 weeks after randomisation.

Primary analysis

1) The primary estimand will be addressed by the below primary analysis based on all subjects included in the FAS and using the in-trial observation period. Note that if subjects withdraw consent to contribute additional information or are completely lost to follow-up, missing data will occur. The primary analysis will be implemented as a statistical model using multiple imputations where the subjects without HbA_{1c} measurements at scheduled visits will have their change from baseline HbA_{1c} value(s) imputed from the available information from the treatment group the subject has been randomised to. Note that this resembles in essence a mixed model of repeated measurements analysis. Subjects without post-randomisation measurements contribute to the analysis, as the missing values will be imputed. The analysis will be implemented as follows:
    - In the first step, intermittent missing values are imputed using a Markov Chain Monte Carlo method, in order to obtain a monotone missing data pattern. This

imputation is done for each group separately and 100 copies of the dataset will be generated.

- In the second step, for each of the 100 copies of the dataset, an analysis of variance model with region and metformin use at baseline (Yes/No) as factors and baseline HbA_{1c} as covariate is fitted to the change in HbA_{1c} from baseline to week 4 for each treatment group separately. The estimated parameters, and their variances, from these models are used to impute missing values at week 4 for subjects in each treatment group, based on region, metformin use at baseline (Yes/No) and baseline HbA_{1c}.

- In the third step, for each of the 100 copies of the dataset, missing values at week 8 are imputed in the same way as for week 4. The imputations are based on an analysis of variance model with region and metformin use at baseline (Yes/No) as factors and baseline HbA_{1c} and change from baseline in HbA_{1c} at week 4 as covariates.

- This stepwise procedure is then repeated sequentially for week 12 and 16

- For each of the complete data sets, the change from baseline to week 16 is analysed using an analysis of variance model with treatment, region and metformin use at baseline (Yes/No) as factors, and baseline HbA_{1c} as a covariate.

- The estimates and standard deviations for the 100 data sets are pooled to one estimate and associated standard deviation using Rubin's formula:

$$m_{MI} = \frac{1}{100}\sum_{i=1}^{100} m_i$$

$$SD_{MI} = \sqrt{\frac{1}{100}\sum_{i=1}^{100} {SD_i}^2 + \left(1+\frac{1}{100}\right)\left(\frac{1}{100-1}\right)\sum_{i=1}^{100}(m_i - m_{MI})^2}$$

  where m_i and SD_i are the estimated means and standard deviations for the 100 copies of the dataset, and m_{MI} and SD_{MI} are the pooled estimates.

- From m_{MI} and SD_{MI}, the 95% confidence interval for the treatment differences is calculated.

Non-inferiority of faster aspart will be considered confirmed if the upper boundary of the two-sided 95% confidence interval is below or equal to 0.4% or equivalent if the p-value for the one-sided test of

$$H_0: D > 0.4\% \quad \text{against } H_A: D \leq 0.4\%,$$

is less than or equal to 2.5%, where D is the mean treatment difference (faster aspart minus NovoRapid®).

Note that as the anticipated number of subjects discontinuing treatment, but not trial is low, multiple imputations based on such subjects is not expected to be suitable.

Provided that the hierarchical testing allows, the evaluation of superiority will be based on the same statistical model, as the primary analysis 1). The associated sensitivity analysis that follows will investigate the robustness of non-inferiority (analyses 3a and 3d) and superiority (analyses 3b and 3c) as well.

Sensitivity analyses for the primary analysis addressing the primary estimand

All sensitivity analyses for the primary analysis addressing the primary estimand will use the in-trial observation period.

2) First the primary analysis in 1) will be repeated, but excluding all factors except from treatment from the multiple imputation and analysis of variance models while still including baseline HbA_{1c} as a covariate. This analysis will explore the influence of the different factors.

3) The primary analysis approach chosen for this trial relies on the assumption that missing data is missing at random (MAR). This assumption implies that the HbA_{1c} for subjects leaving the trial, after their withdrawal, develops in a similar way as the HbA_{1c} for similar subjects that remain in the trial (not necessarily on treatment) and had similar development of HbA_{1c} before withdrawal. The MAR assumption may be questionable for subjects withdrawing at own will. Therefore the statistical model using multiple imputation will be repeated with the following alterations:

    a) Imputations will be done from the treatment arm that the subject was randomised to and a value of 0.4% (non-inferiority margin) is added to the change from baseline in HbA_{1c} at week 16 for subjects randomised to faster aspart with an imputed value at week 16[2]. This will serve as a sensitivity analysis for the non-inferiority analysis.

    b) Imputations will be done from the comparator arm (NovoRapid®). This will serve as a supplementary sensitivity analysis for the superiority analysis. The imputation will be done conditional on observed information for subjects on faster aspart without a measurement at week 16 such that the treatment effect diminishes gradually (copy reference/conditional imputation). It does not rely on the MAR assumption, but

assumes that subjects on faster aspart without a measurement at week 16 switch to NovoRapid®. The analysis will use data from the in-trial observation period.

c) Imputation will be done from the comparator arm (NovoRapid®). This will serve as a supplementary sensitivity analysis for the superiority analysis. The imputation will be done with no regard to observed information for subjects on faster aspart without a measurement at week 16 such that the treatment effect diminishes immediately (jump to reference/unconditional imputation). It does not rely on the MAR assumption, but assumes that subjects on faster aspart without a measurement at week 16 switch to NovoRapid®. The analysis will use data from the in-trial observation period.

d) A tipping point analysis based on a statistical model using multiple imputations similar to 1), using the in-trial observation period, will be made. In this analysis, observations for subjects without a measurement are imputed based on the treatment arm they were randomised to and subjects on faster aspart without a measurement are given a penalty. This is done to investigate the robustness of the conclusion in the primary analysis with respect to the MAR assumption and mimics a scenario where the HbA_{1c} of the subjects without a measurement in the faster aspart group evolve less favourably than predicted. As a first step imputations will be done without penalty assuming MAR in the treatment group. Second, the imputed values for week 16 in the faster aspart group will be added a penalty. This is done repeatedly, gradually increasing the penalty until the conclusion from the non-inferiority analysis no longer holds. This will serve as a sensitivity analysis for the non-inferiority analysis and the specific value of the penalty that changes the conclusion will be used to evaluate the robustness of the conclusion of the non-inferiority analysis.

Analyses addressing the secondary estimand

All analyses addressing the secondary estimand will use the on-treatment observation period.

4) The secondary estimand will be analysed using the same statistical model using multiple imputations as the primary analysis in 1) except using the on-treatment observation period.

5) A tipping point analysis based on a statistical model using multiple imputations, similar to sensitivity analysis 3)-d except using the on-treatment observation period.

6) A tipping point analysis based on a statistical model using multiple imputation, similar to 5) but with the modification that subjects without a measurement that discontinued treatment due to non-eligibility (Subjects discontinuing randomised treatment prematurely due to criteria 1, 2, 3, and 4, which are defined in section 6.6 in the protocol) in the faster aspart

group will not have a penalty added to the imputed values. This analysis is motivated by the fact that data from subjects prematurely discontinuing randomised treatment due to non-eligibility can reasonably be assumed to be missing completely at random.

7) The same statistical model using multiple imputations as the analysis in 4), but using the PP analysis set and analysed using the on-treatment observation period. This analysis will investigate the situation that subjects might have deviated from the inclusion and exclusion criteria and will serve as sensitivity analysis for the non-inferiority analysis.

2.4 Secondary endpoints

2.4.1 Confirmatory secondary endpoints

If the effect of treatment with faster aspart can be confirmed in the primary analysis, the trial also aims to confirm effect of treatment with faster aspart for a number of secondary confirmatory endpoints using a hierarchical (fixed sequence) testing procedure as described in section 2 (General consideration). This is based on a priority ordering of the null-hypotheses, and testing them in this order using the two-sided 95% confidence interval approach until an insignificant result appears. The effect is that rejection of the null hypothesis will only be confirmed for endpoints where all previous null-hypotheses have been rejected in favour of faster aspart.

The confirmatory secondary endpoints are:

- Change from baseline in 1-hour PPG increment 16 weeks after randomisation (meal test)
- Change from baseline in 1,5-anhydroglucitol 16 weeks after randomisation

The steps in the hierarchical testing procedure are as follows:

Step 1 (Primary analysis): HbA_{1c} non-inferiority of faster aspart versus NovoRapid®

Step 2: 1-hour PPG increment (meal test) superiority of faster aspart versus NovoRapid®

Step 3: HbA_{1c} superiority of faster aspart versus NovoRapid®

Step 4: 1,5-anhydroglucitol superiority of faster aspart versus NovoRapid®

The primary estimand for the primary endpoint will be repeated for the confirmatory secondary endpoints, change from baseline in 1-hour PPG increment (meal test) and change from baseline in 1,5-anhydroglucitol 16 weeks after randomisation. The analyses related to these estimands are defined below and will be used for the decisions to continue or not, throughout the hierarchical testing procedure. These analyses will be based on the FAS and use the in-trial observation period.

As sensitivity analysis the secondary analysis 4) will also be repeated for the confirmatory secondary endpoints. The analyses will be based on the FAS and using the on-treatment observation period.

Change from baseline in 1-hour PPG increment 16 weeks after randomisation (meal test) (step 2)

As the second step of the hierarchical testing procedure change from baseline in 1-hour PPG increment (meal test) 16 weeks after randomisation will be tested for superiority of faster aspart compared to Novo Rapid®.

The 1-hour PPG increment will be analysed based on the laboratory measured values in the meal test, and is derived as the 1-hour PPG measurement minus the pre-prandial PG measurement.

The 1-hour PPG increment endpoint will be analysed using the FAS and the in-trial observation period based on a multiple imputation technique where the change from baseline in 1-hour PPG increment at week 16 for subjects with missing value are imputed based on data from subjects in the NovoRapid® arm with non-missing values at week 16. Multiple copies (100 copies) of the full dataset will be generated by imputing missing values based on estimated parameters as follows:

- An analysis of variance model with region and metformin use at baseline (Yes/No) as factors and baseline 1-hour PPG increment as covariate is fitted to the change from baseline in 1-hour PPG increment at week 16 for the NovoRapid® group only. The estimated parameters, and their variances, from this model are used to impute missing values using stochastic simulation at week 16 for subjects in both treatment groups in order to generate 100 complete datasets.
- For each of the complete data sets, the change from baseline to week 16 is analysed using an analysis of variance model with treatment, region and metformin use at baseline (Yes/No) as factors, and baseline 1-hour PPG increment as covariate.
- The estimates and standard deviations for the 100 data sets are pooled to one estimate and associated standard deviation using Rubin's formula. From this, the pooled estimate and 95% confidence interval for the treatment difference are calculated.

The superiority will be assessed by comparing the upper limit of the 95% CI to 0. If the upper 95% CI is below 0 then superiority will be confirmed.

Change from baseline in HbA_{1c} 16 weeks after randomisation (step 3)

Step 3 in the hierarchical testing procedure is to confirm superiority of change from baseline HbA_{1c} 16 weeks after randomisation of the effect of treatment with faster aspart compared to NovoRapid®. Superiority will be addressed based on the same 95% CI that was used for addressing the primary

analysis. Superiority will be confirmed if the upper boundary of the two-sided 95% confidence interval of the mean treatment difference (faster aspart minus NovoRapid®) is below 0.

Change from baseline in 1,5-anhydroglucitol 16 weeks after randomisation (step 4)

Step 4 in the hierarchical testing procedure is to confirm superiority of change from baseline in 1,5-anhydroglucitol 16 weeks after randomisation of the effect of treatment with faster aspart compared to NovoRapid®. The endpoint will be analysed based on all planned post baseline measurements until or at 16 weeks after randomisation using a model similar to the primary analysis 1), but with baseline 1,5-anhydroglucitol as covariate. Superiority will be confirmed if the upper boundary of the two-sided 95% confidence interval of the mean treatment difference (faster aspart minus NovoRapid®) is below 0.

2.4.2 Supportive secondary endpoints

2.4.2.1 Efficacy endpoints

All endpoints except insulin dose in this section will be assessed using the FAS and the in-trial observation period and repeated using the on-treatment observation period. Insulin dose will be presented using the safety analysis set and will therefore only use the on-treatment observation period.

Change from baseline in FPG 16 weeks after randomisation

Change from baseline in FPG 16 weeks after randomisation will be analysed based on all planned post baseline measurements until or at 16 weeks after randomisation using a model similar to 1) except with baseline FPG as covariate.

If a subject achieves HbA_{1c} target 16 weeks after randomisation

HbA_{1c} < 7.0%

A dichotomous (responder/non-responder) endpoint will be defined based on whether a subject has met the HbA_{1c} target (HbA_{1c} <7.0%) 16 weeks after randomisation.

This responder endpoint will be analysed based on a logistic regression model using treatment, region and metformin use at baseline (Yes/No) as factors, and baseline HbA_{1c} as covariate. In analysis of the in-trial observation period subjects without an HbA_{1c} measurement at week 16 will be treated as non-responders. In the analysis using the on-treatment observation period subjects who discontinue randomised treatment, initiate ancillary treatment, withdraw from trial, and/or have no HbA_{1c} measurement at week 16 are included as non-responders.

HbA_{1c} < 7.0 % without severe hypoglycaemia

A dichotomous (responder/non-responder) endpoint will be defined based on whether a subject has met the HbA_{1c} target (HbA_{1c} <7.0%) 16 weeks after randomisation without treatment emergent severe hypoglycaemic episodes.

This responder endpoint will be analysed based on a logistic regression model using treatment, region and metformin use at baseline (Yes/No) as factors and baseline HbA_{1c} as covariate. In the analysis of the in-trial observation period subjects without an HbA_{1c} measurement at week 16 will be treated as non-responders. In the analysis using the on-treatment observation period subjects who discontinue randomised treatment, initiate ancillary treatment, withdraw form trial, and/or have no HbA_{1c} measurement at week 16 will be included as non-responders.

Change from baseline in 30- minutes, 1- hour, 2- hour, 3- hour and 4- hour PPG and 30-minutes, 2- hour, 3- hour and 4- hour PPG increment 16 weeks after randomisation (meal test)

Laboratory measured PG from the meal test will be analysed for 30 minutes, 1-hour, 2-hours, 3-hours and 4-hours PPG separately. The corresponding PPG increments will be derived separately using each PPG measurement minus the pre-prandial PG measurement.

Change from baseline in PPG and PPG increment 16 weeks after randomisation will be analysed separately using a model similar to the model used in hierarchical testing procedure step 2 for 1-hour PPG increment 16 weeks after randomisation (meal test) except with the corresponding baseline value as covariate.

Change from baseline in endpoints derived from the 7-9-7-point SMPG profile 16 weeks after randomisation

In general, analyses will be based on the entire 7-9-7-point SMPG profile except for the analyses of nocturnal endpoints where information in the 9-point SMPG profile will be utilised.

PPG and PPG increments based on the 7-9-7-point SMPG profiles will be derived separately for PG measurements made 1 hour after the meal. In the following section this distinction will be considered implicit and without further explanation.

Pre-prandial PG and PPG will be recorded by the subjects as part of the 7-9-7-point SMPG profile prior to three defined visits. Individual mean mealtime PPG (post-breakfast, post-lunch, post-main evening meal) will be derived from the three profiles. Overall mean PPG will be derived from the individual derived mealtime PPG values.

PPG increment for each meal (breakfast, lunch, main evening meal) will be derived from the 7-9-7-point SMPG profile as the difference between PPG values and pre-prandial PG values in each

separate profile. The mean of the derived increments will then be calculated separately for each meal. Mean 1 hour PPG increments over all meals will be derived as the mean of all corresponding mean meal increments.

- Change from baseline in mean of the 7-9-7-point SMPG profile

  The mean of the 7-9-7-point SMPG profile is defined as the area under the curve profile divided by the measurement time, and is calculated using the linear trapezoidal technique.

  Change from baseline in the mean of the 7-9-7-point SMPG profile 16 weeks after randomisation will be analysed based on all planned post baseline measurements until or at 16 weeks after randomisation using a model similar to 1) except with the corresponding baseline value as covariate.

- Change from baseline in PPG and PPG increment (mean, breakfast, lunch and main evening meal)

  Change from baseline in PPG and PPG increment endpoints 16 weeks after randomisation for mean over all three meals and the individual meals will be analysed separately based on all planned post baseline measurements until or at 16 weeks after randomisation using a model similar to 1) except with the corresponding baseline value as covariate.

- Fluctuation in 7-9-7-point SMPG profile

  The fluctuation in the 7-9-7-point SMPG profile is defined as:

$$\frac{1}{T}\int_0^T |PG(t) - \overline{PG}|dt$$

  where *T*, *PG(t)* and $\overline{PG}$ denotes the length of the profile, the *PG* value at time *t* and the mean of the profile, respectively. It will be calculated using the linear trapezoidal technique.

  Fluctuation in the 7-9-7-point SMPG profile will be logarithmically transformed and analysed based on all planned post baseline measurements until or at 16 weeks after randomisation using a model similar to 1) except with the corresponding log-transformed baseline value as covariate. Estimated treatment means and the estimated treatment differences with corresponding 95% CI will be back-transformed to the original scale, resulting in estimated geometric means, a treatment ratio and a 95% CI for the treatment ratio.

- Change from baseline in nocturnal SMPG measurements

Change from baseline in nocturnal SMPG measurements will be assessed by considering the difference between PG values available at bedtime, at 4 AM and the before breakfast value the following day: (4 AM PG value minus at bedtime PG value), (before breakfast PG value minus at bedtime PG value) and (before breakfast PG value minus 4 AM PG value).

Change from baseline in nocturnal SMPG measurements 16 weeks after randomisation will be analysed based on all planned post baseline measurements until or at 16 weeks after randomisation using a model similar to 1) except with the corresponding baseline values as covariate.

If a subject achieves PPG target (based on overall mean of daily PPG measurements in 7-9-7-point SMPG profile) 16 weeks after randomisation:

Overall PPG (1 hour) ≤7.8 mmol/L [140 mg/dL]

A dichotomous (responder / non-responder) endpoint will be defined based on whether a subject has reached an overall mean 1 hour PPG ≤7.8 mmol/L [140 mg/dL] 16 weeks after randomisation, where PPG is derived from the 7-9-7-point SMPG profile.

This responder endpoint will be analysed based on a logistic regression model using treatment, region and metformin use at baseline (Yes/No) as factors, and baseline overall 1-hour mean PPG as covariate. In analysis of the in-trial observation period subjects without an overall mean 1 hour PPG at week 16 will be treated as non-responder. In the on-treatment observation period analysis subjects who discontinue randomised treatment, initiate ancillary treatment, withdraw from trial, or have no 1-hour mean PPG measurement at week 16 will be included as non-responders.

Overall PPG (1-hour) ≤7.8 mmol/L [140 mg/dL] without severe hypoglycaemia

A dichotomous (responder / non-responder) endpoint will be defined based on whether a subject has reached an overall 1-hour PPG ≤7.8 mmol/L [140 mg/dL] 16 weeks after randomisation without any treatment emergent severe hypoglycaemic episodes.

This responder endpoint will be analysed based on a logistic regression model using treatment, region and metformin use at baseline (Yes/No) as factors, and baseline mean 1-hour PPG as covariate. In analysis of the in-trial observation period subjects without an overall mean 1 hour PPG at week 16 will be treated as non-responders. In the on-treatment observation period analysis subjects who discontinue randomised treatment, initiate ancillary treatment, withdraw from trial, or have no 1-hour mean PPG measurement at week 16 will be included as non-responders.

Insulin dose (Units/day and Units/kg/day; total basal, total bolus, total daily insulin dose and individual meal insulin dose) 16 weeks after randomisation

The insulin doses will be summarised descriptively by treatment week according to regimen, both by meal type and as total daily dose in units and units/kg (total daily and separately for each mealtime dose). Insulin doses will be summarised using the on-treatment observation period and using the safety analysis set.

Lipids-lipoproteins profile 16 weeks after randomisation

Lipid endpoints (total cholesterol, HDL cholesterol, LDL cholesterol) will be logarithmically transformed and analysed separately based on all planned post baseline measurements until or at 16 weeks after randomisation using a model similar to 1) except with the corresponding log-transformed baseline measurement as covariate. Estimated treatment means and the estimated treatment difference with corresponding 95% CI will be back-transformed to the original scale, resulting in estimated geometric means, a treatment ratio and a 95% CI for the treatment ratio.

2.4.2.2 Safety endpoints

In terms of AEs, as a minimum, SAEs will be tabulated separately using the in-trial observation period.

All events in the in-trial observation period will be listed with information about whether it appeared in the on-treatment observation period or not.

Number of treatment emergent adverse events

Adverse events (AEs) will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). AEs will be presented based on system organ class and preferred terms.

A Treatment Emergent Adverse Event (TEAE) is defined as an event that has an onset date on or after the first day of exposure to randomised treatment, and no later than seven days after the last day of exposure to randomised treatment.

TEAEs are summarised descriptively, whereas AEs not defined as treatment emergent are presented in listings, including AEs reported in the 30-day follow-up period. The summaries of TEAEs are made displaying the number of subjects with at least one event, the percentage of subjects with at least one event, the number of events and the event rate per 100 patient years of exposure. These summaries are done by seriousness, severity, relation to insulin treatment, relation to technical complaint, premature treatment discontinuation due to AEs, and outcome.

Furthermore, summary tables based on system organ class and preferred terms are made for:

- All TEAEs

- Serious TEAEs
- Possibly or probably related TEAEs
- Severe TEAEs
- TEAEs with preferred term that are experienced by at least 5% of the subjects in any treatment arm or by at least 5% of all subjects

For AEs where additional information is recorded, this will be listed.

AEs occurring during the run-in period are considered non-treatment emergent and will be summarised separately.

Number of treatment emergent injection site reactions

Treatment emergent injection site reactions occurring during the trial will be summarised and listed.

Classification of Hypoglycaemia

Treatment emergent: hypoglycaemic episodes will be defined as treatment emergent if the onset of the episode occurs on or after the first day of exposure to randomised treatment, and no later than one day after the last day of exposure to randomised treatment.

Nocturnal hypoglycaemic episodes: episodes occurring between 00:01 and 05.59 both inclusive.

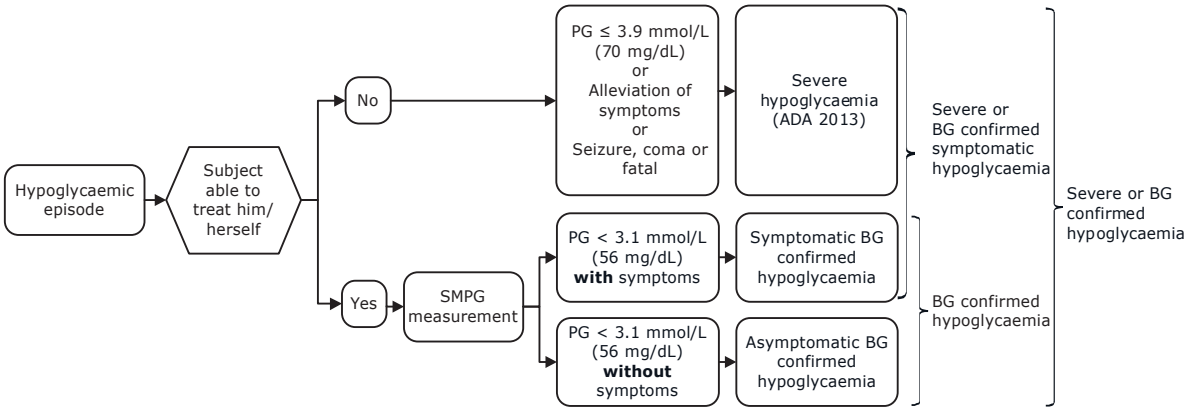
Hypoglycaemic episodes are classified according to the Novo Nordisk classification of hypoglycaemia (see Figure 2–1) and the ADA classification of hypoglycaemia (see Figure 2–2).

Novo Nordisk classification of hypoglycaemia

In normal physiology, symptoms of hypoglycaemia occur below a PG level of 3.1 mmol/L (56 mg/dL).[3] Therefore, Novo Nordisk has included hypoglycaemia with PG levels below this cut-off point in the definition of BG confirmed hypoglycaemia.

Novo Nordisk uses the following classification (see Figure 2–1) in addition to the ADA classification:

- Severe hypoglycaemia according to the ADA classification.[4]
- Severe or BG confirmed hypoglycaemia: An episode that is severe according to the ADA classification[4] or BG confirmed by a PG value <3.1 mmol/L (56 mg/dL) **with** or **without** symptoms consistent with hypoglycaemia.



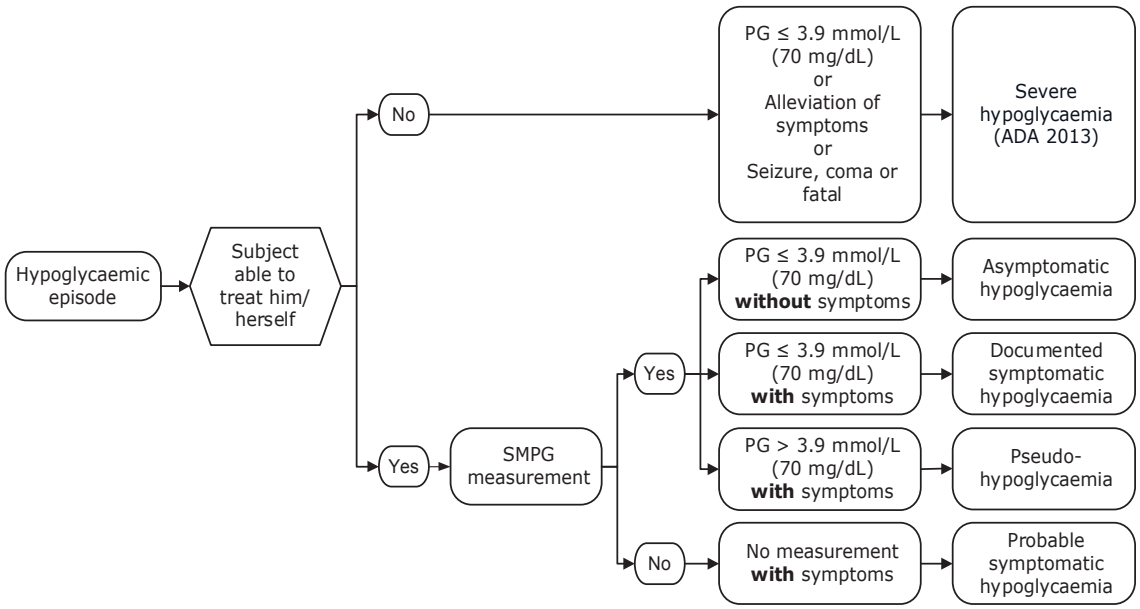
Note: Glucose measurements are performed with capillary blood calibrated to plasma equivalent glucose values

BG: blood glucose PG: plasma glucose SMPG: Self-measured plasma glucose

Figure 2–1 Novo Nordisk classification of hypoglycaemia

ADA classification[4] of hypoglycaemia

- Severe hypoglycaemia: An episode requiring assistance of another person to actively administer carbohydrate, glucagon, or take other corrective actions. Plasma glucose concentrations may not be available during an event, but neurological recovery following the return of PG to normal is considered sufficient evidence that the event was induced by a low PG concentration.
- Asymptomatic hypoglycaemia: An episode not accompanied by typical symptoms of hypoglycaemia, but with a measured PG concentration ≤ 3.9 mmol/L (70 mg/dL).
- Documented symptomatic hypoglycaemia: An episode during which typical symptoms of hypoglycaemia are accompanied by a measured PG concentration ≤ 3.9 mmol/L (70 mg/dL).
- Pseudo-hypoglycaemia: An episode during which the person with diabetes reports any of the typical symptoms of hypoglycaemia with a measured PG concentration > 3.9 mmol/L (70 mg/dL) but approaching that level.
- Probable symptomatic hypoglycaemia: An episode during which symptoms of hypoglycaemia are not accompanied by a PG determination but that was presumably caused by a PG concentration ≤ 3.9 mmol/L (70 mg/dL).



Note: Glucose measurements are performed with capillary blood calibrated to plasma equivalent glucose values

PG: plasma glucose SMPG: Self-measured plasma glucose

Figure 2–2 ADA classification of hypoglycaemia

Treatment emergent hypoglycaemic episodes are presented in terms of the number of subjects with at least one event (N), the percentage of subjects with at least one event (%), the number of events (E) and the event rate per 100 years of exposure (R). Separate summaries are made by severity considering all episodes, nocturnal and daytime episodes using Novo Nordisk and ADA classified episodes. All episodes will also be summarised by category, including summaries in relation to time since start of meal, as occurring within the following time intervals:

- During first 1, 2, and 4 hours after start of meal
- Between 1 (exclusive) to 2 hours (inclusive) after start of meal
- Between 2 (exclusive) to 3 hours (inclusive) after start of meal
- Between 3 (exclusive) to 4 hours (inclusive) after start of meal
- Between 2 (exclusive) to 4 hours (inclusive) after start of meal

The number of treatment emergent severe or BG confirmed hypoglycaemic episodes (all, daytime, nocturnal, 1 hour, 2 hours, 4 hours, 1 (exclusive) to 2 hours (inclusive), 2 (exclusive) to 3 hours (inclusive), 3 (exclusive) to 4 hours (inclusive), and 2 (exclusive) to 4 hours (inclusive) after start of the meal) will be analysed based on the FAS using a negative binomial regression model with a log - link function and the logarithm of the time period in which a hypoglycaemic episode is consid ered

treatment emergent as offset. The model will include treatment, region and metformin use at baseline (Yes/No) as factors. To the extent where data allow, separate analysis will be performed for severe hypoglycaemic episodes (all).

Change from baseline in clinical evaluations 16 weeks after randomisation:

Physical examination

The physical examination parameters (head, ears, eyes, nose, throat, neck, respiratory system, cardiovascular system, gastrointestinal system incl. mouth, musculoskeletal system, central and peripheral nervous system, skin), and their change from baseline, will be summarised descriptively using the on-treatment period. All findings will be listed.

Vital signs

Vital signs include diastolic blood pressure, systolic blood pressure and pulse. The measurements will be summarised descriptively using the on-treatment period and both actual values as mean change and change from baseline.

Electrocardiogram

ECG findings will be summarised descriptively using the on-treatment period and including summaries of the change from baseline. Change from baseline will be summarised as normal/abnormal not clinically significant/abnormal clinically significant categorisation in shift tables.

Fundoscopy/fundus photography

Fundoscopy/fundus photography findings will be summarised descriptively using the on-treatment period and including summaries of the change from baseline. Change from baseline will be summarised as normal/abnormal not clinically significant/abnormal clinically significant categorisation in shift tables.

Change from baseline in clinical laboratory assessments 16 weeks after randomisation

Change from baseline 16 weeks after randomisation in central laboratory assessments:

- Haematology (erythrocytes, haematocrit, haemoglobin, leucocytes, thrombocytes)
- Biochemistry (ALT, AST, albumin, alkaline phosphatase, creatinine, potassium, sodium, total bilirubin, total protein)

Individual laboratory values will be compared to their relevant reference range (when existing) and flagged as being below or above the range. The measurements and their change from baseline will

be summarised descriptively using the on-treatment period. Change from baseline will be summarised both the actual values and the low/normal/high categorisation in shift tables.

Change from baseline in body weight and body mass index 16 weeks after randomisation

The measurements will be summarised descriptively using the on-treatment period and the actual values as mean change.

Change from baseline in body weight will be analysed based on all planned post baseline measurements until or at 16 weeks after randomisation using a statistical model similar to 1) except with the corresponding baseline measurement as covariate. The analysis will be based on the safety analysis set and the on-treatment observation period.

2.5 Supplementary analyses

Discontinuation of the glycaemic data collection system

Following the observation of unusual glycaemic data patterns, the original glycaemic data collection system (combined use of MyGlucoHealth Wireless BG-meter and an electronic diary was discontinued during the trial, and subjects were switched to a new BG-meter and paper diary.

A dichotomous variable will be defined according to whether a subject had experienced more or less than 6 weeks between the date of switch of glycaemic data collection system and date of randomisation. A supplementary analysis will be performed by repeating the primary analysis in 1), but including the aforementioned dichotomous variable as a factor, in addition to the already included factor variables. The analysis will be under the framework of the primary estimand, i.e. based on the full analysis set and the in-trial observation period.

Different meal tests between countries

Subjects in Russia, Serbia and Argentina received a different meal test (volume) compared to subjects in all other countries. A dichotomous variable will be defined according to whether or not a subject comes from one of the aforementioned three countries and so received the non-standard meal test. Supplementary analyses will be carried out by repeating the analyses for change from baseline in 30-minutes, 1-hour, 2-hours, 3-hours and 4-hours PPG increment (meal test) 16 weeks after randomisation, but including the aforementioned dichotomous variable as a factor in addition to the already included factor variables. The analyses will be under the framework of the primary estimand, i.e. based on the full analysis set and the in-trial observation period.

3 Changes to the statistical analyses planned in the protocol

Minor editorial corrections have been made throughout the document.

General considerations

It has been clarified how data collected at visit 30A are handled.

Statistical analysis

It has been clarified that the family-wise type 1 error rate will be controlled under the framework of the primary estimand.

In the primary analysis 1) for change from baseline in HbA_{1c}, it has been clarified that any missing change from baseline HbA_{1c} value will be imputed.

In the sensitivity analysis 2) for change from baseline in HbA_{1c}, it has been clarified that baseline HbA_{1c} is included as a covariate in both the multiple imputation model and analysis of variance model.

The descriptions of analyses for change from baseline in HbA_{1c} and PPG and PPG increment have been revised to clarify that all subjects without a measurement at week 16 (not only subjects withdrawing from trial in general) should have their values imputed. The new description is aligned with the primary analysis.

For the sensitivity analyses for the primary analysis addressing the primary estimand, it has been clarified for each analysis whether it serves as a sensitivity analysis for the non-inferiority analysis or the superiority analysis.

In order to emphasize the difference between sensitivity analyses 3b) and 3c), the wording has been rearranged.

Tipping point analysis 5) has been repeated for the in-trial observation period in the additional sensitivity analysis 3d).

For change from baseline in 1-hour PPG increment (meal test) it has been clarified how the endpoint is derived. The derivation is aligned with the derivation for the other endpoints based on the meal test. It has been clarified that imputation is to be performed based on subjects in the NovoRapid® arm with non-missing values at week 16 instead of completers in the NovoRapid® arm.

In the statistical analyses for subjects reaching HbA_{1c} or PPG (SMPG) targets, it has been clarified that subjects without an HbA_{1c} or 1-hour mean PPG measurement at week 16 will be included as non-responders in analyses of both the in-trial and on-treatment observation periods. To reflect that the definition of the on-treatment observation period takes into account ancillary treatment, it has been clarified that subjects who initiate ancillary treatment are also included as non-responders in analyses of the on-treatment observation period.

"Change from baseline" has been removed for fluctuation in the 7-9-7-point SMPG profile and lipids-lipoprotein profile because data are logarithmically transformed and change from baseline potentially include negative values.

Adverse events

Since data on AEs leading to withdrawal is not collected, this information can not be summarised.

Summary table for possibly "and" probably related TEAEs has been changed to possibly "or" probably related TEAEs in accordance with previous trials.

Hypoglycaemic episodes

The endpoints treatment emergent hypoglycaemic episodes occurring between 1 (exclusive) to 2 hours (inclusive), 2 (exclusive) to 3 hours (inclusive), 3 (exclusive) to 4 hours (inclusive) after start of meal have been added to further investigate the safety of faster aspart.

It has been clarified that a separate statistical analysis of number of treatment emergent severe hypoglycaemic episodes will only be performed on the overall number of episodes, not subdivided into daytime and nocturnal or according to timing after start of meal.

Change from baseline in clinical laboratory assessments 16 weeks after randomisation

Total protein has been added to biochemistry in accordance with section 8.5.2 in the protocol.

Change from baseline in body weight and body mass index 16 weeks after randomisation

In order to align with previous trials, the analysis of change from baseline in body weight will be based on the safety analysis set and the on-treatment observation period.

Supplementary analyses

To account for the potential additional variability introduced by the switch of glycaemic data collection system, a supplementary analysis for the primary endpoint (HbA1c) has been specified.

To account for the potential additional variability introduced by using different meal test in different countries, supplementary analyses addressing the endpoints of PPG increments (meal test) have been specified.

4 References

1. International Conference on Harmonisation. ICH Harmonised Tripartite Guideline E9. Statistical Principles for Clinical Trials. 5 Feb 1998.
2. Koch GG. Comments on 'Current issues in non-inferiority trials' by Thomas R. Fleming, Statistics in Medicine, DOI: 10.1002/sim.2855. Stat Med. 2008;27(3):333-42.
3. Schwartz NS, Clutter WE, Shah SD, Cryer PE. Glycemic thresholds for activation of glucose counterregulatory systems are higher than the threshold for symptoms. J Clin Invest. 1987;79(3):777-81.
4. Seaquist ER, Anderson J, Childs B, Cryer P, Dagogo-Jack S, Fish L, et al. Hypoglycemia and diabetes: a report of a workgroup of the American Diabetes Association and the Endocrine Society. Diabetes Care. 2013;36(5):1384-95.