# Cover Page for Statistical Analysis Plan

| | |
|---|---|
| Sponsor name: | Novo Nordisk A/S |
| NCT number | NCT03687827 |
| Sponsor trial ID: | NN1250-4419 |
| Official title of study: | A randomised, cross-over, open-label, multi-centre trial comparing the effect of insulin degludec and insulin glargine 100U/mL, with or without OADs in subjects with type 2 diabetes using flash glucose monitoring |
| Document date | 04 May 2020 |

### 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: NN1250-4419

**Title:** A randomised, cross-over, open-label, multi-centre trial comparing the effect of insulin degludec and insulin glargine 100U/mL, with or without OADs in subjects with type 2 diabetes using flash glucose monitoring

**SWITCH PRO**

Author

[redacted]

*Biostatistics Insulin and Devices*

This ~~confidential~~ document is the property of Novo Nordisk. ~~No unpublished information contained herein may be disclosed without prior written approval from Novo Nordisk. Access to this document must be restricted to relevant parties.~~

# Table of contents

# List of abbreviations

| | |
|---|---|
| ADA | American Diabetes Association |
| CGM | continuous glucose monitoring |
| FGM | flash glucose monitoring |
| HbA1c | glycated haemoglobin |
| LLOQ | lower limit of quantification |
| OAD(s) | oral anti-diabetic drug(s) |
| OD | once daily |
| PP | per protocol |
| PYE | person years of exposure |
| SAP | statistical analysis plan |
| T2DM | type 2 diabetes mettilus |
| ULOQ | upper limit of quantification |

# 1 Introduction

## 1.1 Trial information

**Objectives, estimand and endpoints:**

**Primary, secondary and explorative objective(s)**

The primary objective is to compare the effect on glycaemic control of IDeg versus IGlar 100U/m in a population of T2DM subjects with or without OADs, using FGM.

**Estimand**

The primary estimand is the treatment difference between IDeg and IGlar 100U/mL with respect to percent time spent in glycaemic target range for all randomised subjects that can use the FGM sensor and adhere to at least 18 weeks of treatment on both insulins. An intercurrent event for the primary estimand include lack of effect from any of the insulins leading to a need for initiation of other diabetes medication(s). As described in section 8.1 in the protocol, treatment intensification may result in trial discontinuation. . The primary estimand will handle these intercurrent events as withdrawals and only use data from completers. The primary estimand is the principle stratum direct effect for the population of completers.

**Primary endpoint**

Percentage of time spent in glycaemic target range 70-180 mg/dL (3.9–10.0 mmol/L) both inclusive, during the 2-week maintenance period using FGM (Visit 19-21 (week 16-17) and Visit 37-39 (week 34-35)).

**Supportive secondary endpoints**

- Percentage of time spent in tight glycaemic target range 70-140 mg/dL (3.9–7.8 mmol/L) both inclusive, during the 2-week maintenance periods using FGM (Visit 19-21 (week 16-17) and Visit 37-39 (week 34-35)).

- Percentage of time spent in nocturnal glycaemic target range 70-140 mg/dL (3.9-7.8 mmol/L) both inclusive, in the nocturnal period (00:01 am – 05:59 am both inclusive) during the 2-week maintenance periods using FGM (Visit 19-21 (week 16-17) and Visit 37-39 (week 34-35)).

- HbA1c (% and mmol/mol) after two weeks of maintenance periods (Visit 19-21 (week 16-17) and Visit 37-39 (week 34-35)).

- Mean glucose levels (mg/dL and mmol/L) during the 2-week maintenance periods using FGM (Visit 19-21 (week 16-17) and Visit 37-39 (week 34-35)).

**Exploratory endpoints**

- Glycaemic variability calculated as standard deviation (mg/dL and mmol/L) and coefficient of variation (% ) of glucose levels during the 2-week maintenance treatment using FGM(Visit 19-21 (week 16-17) and Visit 37-39 (week 34-35)).

- Mean insulin dose (IU) during the 2-week maintenance treatment period (Visit 19-21 (week 16-17) and Visit 37-39 (week 34-35)).

- Percentage of time spent in the hypoglycaemia alert range (Level 1) 54-69 mg/dL (3.0-3.8 mmol/L) both inclusive, during the 2-week maintenance periods using FGM (Visit 19-21 (week 16-17) and Visit 37-39 (week 34-35)).

- Percentage of time spent in the clinically significant hypoglycaemic range (Level 2) <54mg/dL (<3.0 mmol/L) during the 2-week maintenance periods using FGM (Visit 19-21 (week 16-17) and Visit 37-39 (week 34-35)).

- Number of clinically significant hypoglycaemic episodes (Level 2) <54mg/dL (<3.0 mmol/L) for at least 15 minutes during the 2-week maintenance periods using FGM (Visit 19-21 (week 16-17) and Visit 37-39 (week 34-35)).

- Number of clinically significant hypoglycaemic episodes (Level 2) <54mg/dL (<3.0 mmol/L) for at least 15 minutes in the nocturnal period (00:01-05:59 AM, both inclusive) during the 2-week maintenance periods using FGM (Visit 19-21 (week 16-17) and Visit 37-39 (week 34-35)).

**Trial design:**

This trial is a 41-week, randomised, cross-over, open-label, multi-center, active controlled trial comparing the effect of insulin IDeg versus insulin IGlar 100U/mL, with or without OADs in subjects with type 2 diabetes using flash glucose monitoring.

The trial includes a screening visit (Visit 1), followed by a 2-week run-in period for eligibility assessment in regards to adherence to FGM requirements. At randomisation (Visit 3), eligible subjects will be randomised 1:1 into one of two treatment periods receiving either IDeg or IGlar 100U/mL. Each treatment period consist of a 16-week titration period followed by a 2-weeks maintenance period where subjects will wear the FreeStyle Libre Pro Sensor. After the first titration and maintenance period the subject will switch to the other trial product, with dose adjustment according to local label, performing the same assessments as in the first treatment period (16-week titration period followed by a 2-weeks maintenance period).

The follow-up period is 1 week.

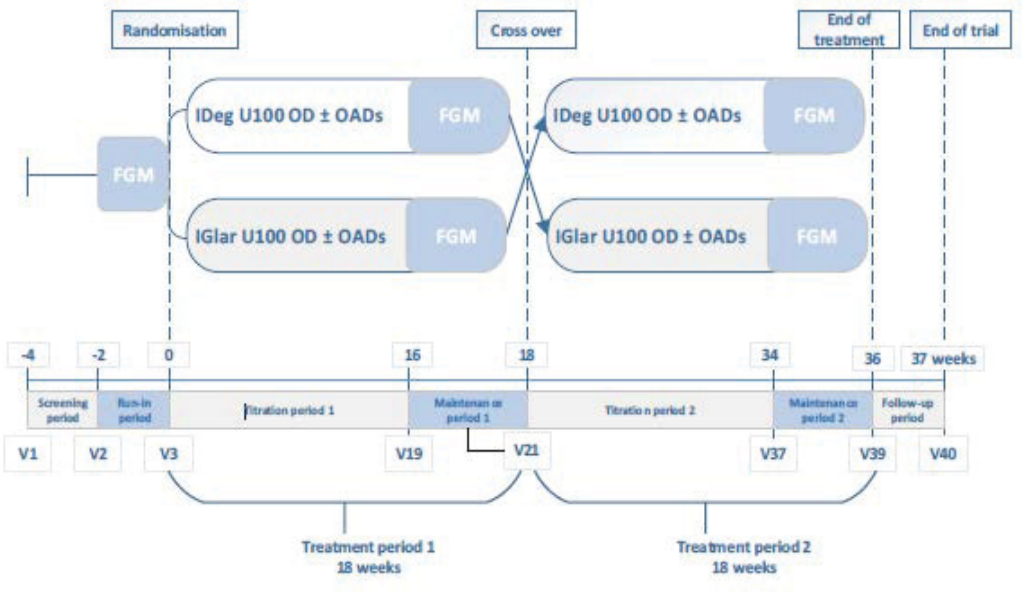


**Figure 1 Trial Design**

**Further details**

For further details on trial design and endpoints please see the NN1250-4419 protocol version 3.0.

### 1.2 Scope of the statistical analysis plan

This SAP is based on the NN1250-4419 protocol, version 3.0.

# 2 Statistical considerations

### 2.1 Sample size determination

The sample size calculation is based on the primary objective for the primary estimand and the primary endpoint of the trial, and is performed as a paired-sample calculation in terms of the mean and standard deviation of the within pair difference within the PP using the paired t-test. Only one previous Novo Nordisk trial using CGM has had a similar design: NN1250-3874 comparing IDeg and IGlar 100U/mL in 24 T1DM subjects using a cross-over design. Regarding time in glycaemic target range (70-180 mg/dL) during the full 24 hours of the day, the within pair difference between IDeg and IGlar 100U/mL had a mean of 0.39 hours and a standard deviation of 2.3. Two previous Novo Nordisk studies have used CGM in T2DM subjects in parallel arm designs, NN1250-3579 and NN1250-3668. However, CGM technology used in those two T2DM trials is outdated compared to the devices used in the current trial, and they are therefore deemed irrelevant as to

guidance on expected difference between the two treatments in this trial. The results from the NN1250-3874 trial are used for sample size calculation despite the differences from the present trial, i.e. indication and length of maintenance treatment. A conservative standard deviation of 2.6 is used. To detect a mean difference of 0.39 with 85% power with a standard deviation of 2.6 based on a paired t-test and 1:1 randomisation requires 401 subjects included in the PP. See Table 3 for sample sizes based on a range of standard deviation and for mean differences of 0.39 and 0.5

| | Effect | |
|---|---|---|
| **SD** | 0.39 | 0.50 |
| 2.0 | 239 | 146 |
| 2.3 | 315 | 192 |
| 2.6 | 401 | 245 |

Sample size is computed for 1:1 randomisation and 85% power. SD: standard deviation.

respectively.

**Table 1 Sample size and power for various standard deviations and effects**

## 2.2 Definition of analysis sets

The full analysis set is all subjects randomised at Visit 3. The safety analysis set is all subjects that are randomised and treated with at least one dose of trial drug after randomisation.

The PP analysis set consists of subjects that stay on assigned treatment until and complete FGM assessment at visits 21 and 39 respectively. Complete FGM assessment is specified as subjects having at least 70% of two weeks FGM measurements per maintenance period (2 * 960 measurements). FGM measurements from the first 24 hours of each fitted sensor are disregarded. All endpoint analysis will be performed on the PP analysis set.

## 2.3 Primary endpoint

The primary endpoint is the percentage of time spent in glycaemic target range (70-180 mg/dL (3.9-10.0 mmol/L), both inclusive) during the 2-week maintenance period. The FreeStyle Libre Pro device stores four measurements per hour and a subject with complete data will contribute with 14*24*4=1344 measurements from each 2-week period. The analysis will discard data from the first 24 hours as per communication with the vendor and so subjects with complete data will contribute with 13*24*4=1248 glucose recordings. The number of recorded measurements may be smaller for various reasons for subjects in the PP analysis set. Following ADA criteria we will demand that at least 70% of two weeks FGM measurements, corresponding to 10 days of FGM data, must be available for subjects to be included in the analysis. The percentage of time spent in glycaemic target range will be calculated as the number of recorded measurements in glycaemic

target range (70-180 mg/dL (3.9-10.0 mmol/L), both inclusive) divided by the total number of recorded measurements.

The primary estimand will be estimated by the primary endpoint analysed in a linear mixed model with treatment and period as fixed effects, subject as a random effect and with an unstructured residual variance matrix.

Using the unblinded data and before DBL, visual tools will be used to determine whether the data, including the primary endpoint, are normally distributed. If it is then found that the data are not normal, an appropriate transformation will be performed and this will be documented in a SAP or the DBL minutes.

Superiority of IDeg OD compared to IGlar 100U/mL OD with respect to percentage of time spent in glycaemic target range will be considered confirmed if non-inferiority is confirmed and the lower limit of the two-sided 95% confidence interval of the difference (IDeg OD - IGlar 100U/mL OD) is entirely above zero or equivalently if the p-value for the one-sided test of

H0: $D \leq 0\%$ against HA: $D > 0\%$

is smaller than 2.5%, where D is the estimated mean difference of time spent in glycaemic target range (IDeg OD - IGlar 100U/mL OD). As such, a fixed hierarchical testing procedure will be applied, which will effectively adjust for multiple testing.

Non-inferiority will be evaluated by comparing the lower limit of the two-sided 95% confidence interval for the difference between IDeg and IGlar 100U/mL to a non-inferiority margin of 0.2h/24h (0.83%). If the lower CI limit is above -0.83%, non-inferiority will be considered established. The two-tailed p-value will be reported for the test of superiority.

Missing data due to intercurrent events such as trial discontinuation is not relevant since the primary estimand is based on the PP analysis set.

A sensitivity analysis will be performed to account for heteroscedasticity due to a varying number of FGM readings within the 70% demand. Specifically, the primary endpoint will be analysed by applying the weight statement in the proc mixed procedure with weights calculated as the square root of the total number of FGM readings.

An additional sensitivity analysis will be performed where all subjects who have FGM readings for at least seven consecutive days will be included. Weighting will be performed as described previously.

Following specifications from the vendor the accuracy of the FGM measurements can be influenced by intake of either ascorbic or/and salicylic acid [1]. The level of consequence depends on the amount of the interfering substance active in the body [1]. Due to the trial design, doses of non-

diabetic concomitant medications are not collected, and it is hereby not possible to identify high risk subjects. A sensitivity analysis will be performed where all FGM measurements, for subjects in the PP analysis set, coinciding with intake of ascorbic and/or salicylic acid, will be excluded. Taking washout periods of both ascorbic and salicylic acid (also applicable for multivitamins including ascorbic acid) into consideration, FGM measurements overlapping the start-to-end dates of intake of either drug + 1 day will be omitted in the analysis. This will result in partial or complete FGM profiles being disregarded. Weighting will be applied as described previously to account for heteroscedasticity due to a varying number of FGM readings

## 2.4 Secondary endpoints

Supportive secondary endpoints

Percentage of time spent in tight glycaemic target range (70-140 mg/dL (3.9-7.8 mmol/L), both inclusive is calculated in the same way as the primary endpoint using the relevant ranges. A sensitivity analysis using weights to account for heteroscedasticity due to a varying number of FGM readings will be performed as described previously.

Percentage of time spent in nocturnal glycaemic target range 70-140 mg/dL (3.9-7.8 mmol/L) both inclusive, in the nocturnal period (00:01 am–05:59 am both inclusive) will be calculated in the same way as the primary endpoint but only using readings with a time stamp in the relevant time period. A sensitivity analysis using weights to account for heteroscedasticity due to a varying number of FGM readings will be performed. As described previously the weights are calculated based on the total number of FGM readings and will in so doing not be adjusted to account for the lower number of nocturnal readings.

All the supportive secondary endpoints regarding time spent in target range are analysed and tested in the same model as the primary endpoint, i.e. one sided tests of the mean difference favouring IDeg must be statistically significantly different from zero at the 0.025 level to support superiority.

As stated in Section 9.1.1.4 of the protocol, sites are instructed not to initiate FGM monitoring periods that will contain daylight saving clock shifts. During data analysis it will be checked if this rule is obeyed, and the length of both the entire day in question in general, and the nocturnal period in question in particular will be adjusted equivalently when it is not. For example, if a subject is under FGM monitoring in the US on 11-MAR-2018 the nocturnal period will start at 00:01AM and end at 04:59 (nominal sensor time) instead of 05:59. In the data analysis one hour will be added to the recorded time points from 02:00 and onwards. In practise then, 11-MAR-2018 will consist of no more than 23h of FGM readings, and the nocturnal period of no more than 4h and 58 minutes of FGM readings. If on the other hand a subject is under FGM monitoring on 04-NOV-2018 in the US one hour will be deducted from the time recorded time points. This will then consist of no more than 25 hours of FGM readings, and the nocturnal period of no more than 6h and 58 minutes of FGM readings. The hour between 01:00-02:00 on 04-NOV will be represented twice in the data,

and to separate them an auxiliary variable stating if a time reading is in the daylight saving period or not will be added to all FGM readings.

The mean glucose levels (mmol/L) will be calculated as the mean of the FGM glucose levels and analysed by models and test equivalent to the primary endpoint. A sensitivity analysis using weights to account for heteroscedasticity due to a varying number of FGM readings will be performed as described previously.

FGM readings have a lower limit of quantification (LLOQ) of 40mg/dL or 2.2 mmol/L and a upper limit of quantification (ULOQ) 500 mg/dL or 27.8 mmol/L. LLOQ marked readings will be assigned the LLOQ value and ULOQ marked readings will be assigned the ULOQ value.

Mean HbA1c values will be analysed using a similar model as compared to the primary endpoint, but only a single two-sided test for statistically significant differences between IDeg and IGlar 100U/mL will be performed.

## 2.5 Exploratory endpoints

Glycaemic variability will be calculated as the standard deviation (mmol/L) and coefficient of variation (%) of the FGM glucose levels during the 2-week maintenance treatment periods. Both will be analysed and tested with models and tests equivalent to the primary endpoint.

Mean insulin dose in international units is calculated as the mean dose over the maintenance treatment periods for each subject. It will be analysed and tested using models and tests equivalent to the primary endpoint.

Percentage of time spent in the hypoglycaemic alert range (54-69 mg/dL (3.0-3.8 mmol/l), both inclusive) and percentage of time spent in the clinically significant hypoglycaemic range (<54 mg/dL (<3.0 mmol/L)) are calculated in the same way as the primary endpoint using the relevant ranges. A sensitivity analysis using weights to account for heteroscedasticity due to a varying number of FGM readings will be performed as described previously.

All the exploratory endpoints regarding time in range are analysed and tested in the same model as the primary endpoint, i.e. one sided tests of the mean difference favouring IDeg must be statistically significantly different from zero at the 0.025 level to support superiority.

Number of hypoglycaemic episodes <54 mg/dL (<3.0 mmol/L) for at least 15 minutes and number of hypoglycaemic episodes <54mg/dL (<3.0 mmol/L) for at least 15 minutes in the nocturnal period (00:01-05:59 AM, both inclusive) will be counted as follows: At least two consecutive FGM glucose readings (i.e., readings that follow each other continuously in time) in the specified range are counted as a hypoglycaemic episode. Two hypoglycaemic episodes must be separated by at least 15 minutes above the hypoglycaemic range, equivalent to 2 consecutive FGM readings above the relevant hypoglycaemic range to be counted as two or more episodes instead of one. For the

nocturnal episodes, only FGM readings in the relevant period will be used. Person years of exposure (PYE) will be calculated as

PYE=(N_FGM*15 minutes)/(365.25 days/year*(24 hours)/day*(60 minutes)/hour)

where N_FGM is the total number FGM readings in the relevant period. The number of hypoglycaemic episodes will be analysed using a negative binomial regression with treatment and period as fixed effects, subject as a normally distributed random effect and the log(PYE) as an offset. A p-value smaller than 0.025 for the one sided test of the rate ratio in favour of IDeg will be considered as support for superiority of IDeg over IGlar 100U/mL. If this model does not converge, a Poisson regression model will be fitted.

# 3 Changes to the statistical analyses planned in the protocol

Updates to section 10 in protocol version 3.0 is described below.

## 3.1 Contribution to per protocol analysis set

p. 38 in section 10.2:

*"The PP analysis set consists of subjects that stay on assigned treatment until and complete FGM assessment at visits 21 and 39 respectively."*

Is replaced by:

*"The PP analysis set consists of subjects that stay on assigned treatment until and complete FGM assessment at visits 21 and 39 respectively. Complete FGM assessment is specified as subjects having at least 70% of two weeks FGM measurements per maintenance period (2 * 960 measurements). FGM measurements from the first 24 hours of each fitted sensor are disregarded. All endpoint analysis will be performed on the PP analysis set.*

Rationale: To ensure unambiguous definition of per protocol analysis set.

## 3.2 Presentation of test hierarchy

p. 39 in section 10.3.1 after:

*"If the lower CI limit is above -0.83%, non-inferiority will be considered established. "*

Is added:

"*The two-tailed p-value will be reported for the test of superiority."*

Rationale: Clarification to ensure legible outputs.

## 3.3 Specification of weights for primary endpoint

p 39 in section 10.3.1:

*"The primary endpoint will be transformed using inverse variance weighting. Specifically, the primary endpoint will be multiplied by a weight calculated as the inverse of the square root of the total number of FGM readings."*

Is replaced by:

*"Specifically, the primary endpoint will be analysed by applying the weight statement in the proc mixed procedure with weights calculated as the square root of the total number of FGM readings.*"

Rationale: Ensuring that full FGM profiles have more impact compared to partial profiles.

### 3.4 Further specification of weights for primary endpoint

p 40 in section 10.3.1:

*"Inverse variance weighting will be performed as described previously."*

Is changed to:

*"Weighting will be performed as described previously."*

Rationale: Due to the change in 3.3.

### 3.5 Sensitivity analysis for primary endpoint

p. 40 section 10.3.1 after:

*"An additional sensitivity analysis will be performed where all subjects who have FGM readings for at least seven consecutive days will be included. Inverse variance weighting will be performed as described previously."*

Is added:

*"Following specifications from the vendor the accuracy of the FGM measurements can be influenced by intake of either ascorbic or/and salicylic acid* [1]*. The level of consequence depends on the amount of the interfering substance active in the body* [1]*. Due to the trial design, doses of non-diabetic concomitant medications are not collected, and it is hereby not possible to identify high risk subjects. A sensitivity analysis will be performed where all FGM measurements, for subjects in the PP analysis set, coinciding with intake of ascorbic and/or salicylic acid, will be excluded. Taking washout periods of both ascorbic and salicylic acid (also applicable for multivitamins including ascorbic acid) into consideration, FGM measurements overlapping the start-to-end dates of intake of either drug + 1 day will be omitted in the analysis. This will result in partial or complete FGM profiles being disregarded. Weighting will be applied as described previously to account for heteroscedasticity due to a varying number of FGM readings"*

Rationale: An additional sensitivity analysis is needed to evaluate the impact of intake of either ascorbic or/and salicylic acid on the accuracy of the FGM measurements during maintenance periods.

### 3.6 Specification of weights for secondary endpoint (Percentage of time spent in tight glycaemic target range)

p. 40 section 10.3.2:

*"A sensitivity analysis using the inverse of the square root of the number of FGM readings as inverse variance weights to account for heteroscedasticity due to a varying number of FGM readings will be performed."*

Is replaced by:

*"A sensitivity analysis using weights to account for heteroscedasticity due to a varying number of FGM readings will be performed as described previously."*

Rationale: Due to the change in 3.3.

### 3.7 Further specification of weights for secondary endpoint (Percentage of time spent in nocturnal glycaemic target range)

p. 40 section 10.3.2:

*"A sensitivity analysis using the inverse of the square root of the number of FGM readings in the relevant time period as inverse variance weights to account for heteroscedasticity due to a varying number of FGM readings will be performed.*"

Is replaced by:

*"A sensitivity analysis using weights to account for heteroscedasticity due to a varying number of FGM readings will be performed. As described previously the weights are calculated based on the total number of FGM readings and will in so doing not be adjusted to account for the lower number of nocturnal readings."*

Rationale: Due to the change in 3.3.

### 3.8 Clarification of secondary endpoint

p. 40 section 10.3.2:

*"The mean glucose levels (mg/dL and mmol/L) will be calculated as the mean of the FGM glucose levels and analysed by models and test equivalent to the primary endpoint. "*

Is reduced to:

*"The mean glucose levels (mmol/L) will be calculated as the mean of the FGM glucose levels and analysed by models and test equivalent to the primary endpoint. "*

Rationale: Mean glucose levels in mg/dL will not be presented in the clinical trial report.

### 3.9 Further specification of weights for secondary endpoint (mean glucose levels)

p. 41 section 10.3.2:

*"A sensitivity analysis using the inverse of the square root of the total number of FGM readings as inverse variance weights will be performed to account for heteroscedasticity due to a varying number of readings pr subject pr. period."*

Is replaced by:

*"A sensitivity analysis using weights to account for heteroscedasticity due to a varying number of FGM readings will be performed as described previously."*

Rationale: Due to the change in 3.3.

### 3.10 Clarification of ULOQ and LLOQ in FGM data

p. 41 section 10.3.2. after:

*"The mean glucose levels (mg/dL and mmol/L) will be calculated as the mean of the FGM glucose levels and analysed by models and test equivalent to the primary endpoint. A sensitivity analysis using the inverse of the square root of the total number of FGM readings as inverse variance weights will be performed to account for heteroscedasticity due to a varying number of readings pr subject pr. period."*

Is added:

*"FGM readings have a lower limit of quantification (LLOQ) of 40mg/dL or 2.2 mmol/L and a upper limit of quantification (ULOQ) 500 mg/dL or 27.8 mmol/L. LLOQ marked readings will be assigned the LLOQ value and ULOQ marked readings will be assigned the ULOQ value."*

Rationale: To inhibit the impact of project standard values on secondary (mean) and exploratory (standard deviation) endpoints.

### 3.11 Clarification of exploratory endpoint

p. 41 section 10.3.3:

*"Glycaemic variability will be calculated as the standard deviation (mg/dL and mmol/L) and coefficient of variation (%) of the FGM glucose levels during the 2-week maintenance treatment periods."*

Is reduced to:

*"Glycaemic variability will be calculated as the standard deviation (mmol/L) and coefficient of variation (%) of the FGM glucose levels during the 2-week maintenance treatment periods."*

Rationale: Glycaemic variability calculated as the standard deviation in mg/dL will not be presented in the clinical trial report.

### 3.12 Specification of weights for exploratory endpoint (Percentage of time spent in the hypoglycaemic alert range)

p. 41 section 10.3.3:

*"A sensitivity analysis using the inverse of the square root of the number of FGM readings as inverse variance weights to account for heteroscedasticity due to a varying number of FGM readings will be performed."*

Is replaced by:

*"A sensitivity analysis using weights to account for heteroscedasticity due to a varying number of FGM readings will be performed as described previously."*

Rationale: Due to the change in 3.3.

# 4 References

[1] I. Abbott Laboratories, “FreeStyle Libre 14 Day Indications and Important Safety Information,” [Online]. Available: https://provider.myfreestyle.com/safety-information.html#FSLsafetyinfo. [Accessed 2 January 2020].