Statistical Analysis Plan

Full title of the Trial: Statistical Analysis Plan for the Flash-glucose monitoring in sub-optimally controlled type 1 diabetes (FLASH-UK): an open-label, multi-centre, randomised, parallel design study to assess the efficacy of flash glucose monitoring in adults with sub-optimally controlled type 1 diabetes

Acronym: FLASH-UK

Trial registration number: NCT03815006 (Clinicaltrials.gov)

SAP version number with dates: Version 1.0, 05/10/2021

Protocol version number and date: This document has been written based on information contained in the study protocol version 4.0, dated 29 Jun 2020

SAP revisions:

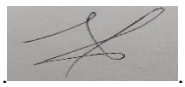
Protocol version	Updated SAP version number	Section number changed	Description of and reason for change	Date changed	Initials

Roles and responsibility: *names, affiliations and roles of SAP contributors*

Ashma Krishan was the original SAP author who was the main contributor to the SAP draft until the SAP was sent to the TMG in October 2020. Vicky Taxiarchi, who took over the role of Trial Statistician in November 2020, is now named as the Lead SAP author and is the below signatory to the SAP.

Roles and responsibility: Signatures

Trial Statistician:

Signature:  .....

Date: .25/Oct/2021
ddmmmyyyy

Name: (please print):
..Vicky Taxiarchi.....

Lead Statistician:

Signature: 

Date: 01/Nov/2021
ddmmmyyyy

Name: (please print):
...Chris Sutton.....

Trial Steering Committee Chair:

Signature: 

Date: 03/Nov/2021
ddmmmyyyy

Name: (please print):
Professor Rory McCrimmon

Chief Investigator:

Signature: .....

Date: ..
01/Nov/2021..
ddmmmyyyy

Name: (please print):
.....Lalantha
Leelarathna.....

Table of Contents

GLOSSARY

AE	Adverse Events
ANCOVA	Analysis of Covariance
CGM	Continuous glucose monitors
CONSORT	Consolidated Standards of Reporting Trials
CSII	Continuous Subcutaneous Insulin Infusion
DEPS-R	The revised Diabetes Eating Problem Survey
D-FISQ	Diabetes fear of injecting and self-testing questionnaire
DTSQ	Diabetes Treatment Satisfaction Questionnaire
FSL	FreeStyle Libre
FSL2	FreeStyle Libre 2
GMSS	The Glucose Monitoring Satisfaction Survey
HbA1c	Glycated haemoglobin (A1c)
IDMC	Independent Data Monitoring Committee
ITT	Intention to Treat
MDI	Multiple daily injections
MICE	Multiple Imputation by Chained Equations
PHQ-9	Patient Health Questionnaire
T1D	Type 1 diabetes mellitus
T1-DDS	Type 1 Diabetes Distress Scale

1 INTRODUCTION

1.1 Background and rationale

Type 1 diabetes mellitus (T1D) is characterised by an absolute deficiency of insulin caused by immunologically-mediated damage to the beta cells in the pancreas and raised blood glucose levels. It is one of the commonest endocrine and metabolic conditions in both children and adults. It is estimated that approximately 415 million adults (5-15% type 1 diabetes) and 520,000 children (95% type 1 diabetes) worldwide suffer from diabetes. In England less than one third of patients with type 1 diabetes achieve a HbA1c level <7.5%. Studies have shown strong relationship with number of finger-stick glucose tests and HbA1c. In contrast to finger-stick glucose monitoring, continuous glucose monitors (CGM) can provide continuous real-time glucose information as well as glucose trend information. However, widespread adoption of these devices has been hampered by several factors including cost, accuracy of earlier devices and user acceptability. In 2014 a new category of device was born: the FreeStyle Libre Flash Glucose Monitoring System (FSL) (Abbott Diabetes Care, Oxon, UK). This device is different to earlier CGM systems. Although it does produce real-time on-demand continuous glucose data, it does not alarm to alert users of rising or falling glucose levels. FreeStyle Libre 2 (FSL2) (which is CE marked) has been produced by the manufacturer. This is identical to FSL but with the optional additional functionality of alarm alerts for users who fall outside of adequately controlled glucose levels.

Use of FSL device in people with well-controlled T1D has shown reduction in hypoglycemia burden. However, to date no randomised study with FSL2 has been undertaken in people with T1D and high HbA1c. The purpose of this study is to determine whether use of flash glucose monitoring with FSL2 device will improve HbA1c over a 24-week randomised period compared to self-monitoring of blood glucose in adults with sub-optimally controlled type 1 diabetes.

1.2 Objectives

1.2.1 Research hypothesis

The null hypothesis is that there is no difference in HbA1c levels over a 24-week period between flash glucose monitoring with FSL2 device and self-monitoring of blood glucose. The alternative hypothesis is whether the HbA1c levels improve with flash monitoring with FSL2 device compared to self-monitoring of blood glucose groups in adults with sub-optimally controlled type 1 diabetes.

1.2.2 Study Objectives

The primary objective is to assess the clinical efficacy of flash glucose monitoring with FSL2 device relative to that with self-monitoring of blood glucose on glycated haemoglobin A1c (HbA1c).

Secondary clinical and psychosocial objectives are, respectively:

- to assess the clinical efficacy of flash glucose monitoring with FSL2 device relative to that with self-monitoring of blood glucose on sensor-based glucose metrics (e.g. time spent in target glucose range 3.9 to 10 mmol/L).
- to evaluate the participants' responses in terms of quality of life, diabetes related distress, diabetes treatment satisfaction, low mood, needle burden and disordered eating behaviours using validated questionnaires.

The safety objective is to evaluate time spent in hypoglycaemia (sensor glucose levels < 3.0 mmol/l and other sensor based biochemical hypoglycaemia) and number of episodes of severe hypoglycaemia with FSL2 and self-monitoring of blood glucose.

2 TRIAL METHODS

2.1 Trial design

An open-label, multi-centre, randomised, parallel study, in adults and adolescents (16 years and older) with type 1 diabetes and sub-optimal glycaemic control (HbA1c 7.5% to 11%), either on insulin pump treatment or multiple daily injections, contrasting flash glucose monitoring using FSL2 device with traditional finger-stick glucose monitoring for 24 weeks.

2.2 Randomisation details

Randomisation to one of the two intervention arms (24-weeks use of flash-glucose monitoring or 24 weeks use of conventional finger-stick glucose monitoring) will use the minimisation method, with a random element to improve allocation concealment. We will minimise over the following factors: study centre (Birmingham; Cambridge; Derby; Manchester; Norwich; Portsmouth), baseline HbA1c (7.5%-9.0%; >9.0%-11%), treatment modality (Multiple daily injections (MDI); Continuous Subcutaneous insulin Infusion (CSII)), prior participation in structured education course (yes; no) and current use of bolus calculator (yes; no).

2.3 Sample size

Assuming a standard deviation of 0.8% and treatment difference of 0.4% - 128 participants (64 per each arm) with primary outcome will give 80% power to detect the difference between treatment groups at 2-sided type 1 error = 5%. Recruitment target is 180 participants (25 to 30 per centre) aiming for 150 to be randomised following the adherence run-in period, and allowing for 15% post-randomisation attrition.

2.4 Framework

The primary and secondary outcomes will test for superiority of the FSL2 device compared to self-monitoring of blood glucose.

2.5 Statistical interim analysis and stopping guidance

2.5.1 Interim Analysis

No formal interim analyses of outcome data will be performed. The study may be stopped if three consecutive participants withdraw on safety grounds or on the advice of an Independent Data Monitoring Committee (IDMC). No statistical early stopping criteria will be implemented.

2.6 Timing of final analysis

All outcome analyses will be undertaken after all baseline, outcome and process data have been entered into the database and the full database is cleaned and locked, and the corresponding checks have been made to the device download data. The Trial Statistician will not perform the final analysis blind to treatment group.

2.7 Timing of outcome assessments

Table 1: Schedule of study visits when participant is randomised to flash glucose monitoring intervention (intervention group).

Visit/contact	Flash glucose monitoring - description	Time since randomisation	Start relative to previous/ next Visit / Activity*
Visit 1	Recruitment & Screening visit: Consent HbA1c, baseline bloods, questionnaires	-2 to -3 weeks	-
Visit 2	Blinded flash glucose monitor insertion	-2 weeks	Within 1 to 2 weeks of Visit 1. Can coincide with Visit 1
Visit 3	Adherence assessment & Randomisation Flash-glucose monitoring initiation - Training, education & competency assessment	0 weeks	After 2 weeks of Visit 2
Visit 4	Review data /optimisation and use of study devices. Data download & collect participant diary	+4 weeks	After 4 weeks of Visit 3
Visit 5	Review data /optimisation. Data download - HbA1c. - Collect participant diary	+12 weeks	After 8 weeks of Visit 4
Visit 6	Not applicable in this arm	+22 weeks	-
Visit 7	End of Flash-glucose monitoring intervention arm - HbA1c - Questionnaires - Data download - Collect participant diary	+24 weeks	12 weeks after Visit 5

* Each study visit can be scheduled with +/- 2 weeks of the planned visit date.

Table 2: Schedule of study visits when participant is randomised self-monitoring of blood glucose

Visit/contact	Self-monitoring of blood glucose - description	Time since randomisation	Start relative to previous / next Visit / Activity*
Visit 1	Recruitment & Screening visit: Consent HbA1c, baseline bloods, questionnaires	-2 to -3 weeks	-
Visit 2	Blinded flash glucose monitor insertion	-2 weeks	Within 1 to 2 weeks of Visit 1. Can coincide with Visit 1
Visit 3	Adherence assessment & Randomisation Self- monitoring of glucose initiation - Education Data download	0 weeks	After 2 weeks of Visit 2
Visit 4	Review data /optimisation Collect participant diary	+4 weeks	After 4 weeks of Visit 3
Visit 5	Review data /optimisation. - HbA1c - Collect participant diary Data download**	+12 weeks	After 8 weeks of Visit 4
Visit 6	Blinded flash glucose monitor insertion (Extra visit in this arm)	+22 weeks	After 10 weeks of Visit 5
Visit 7	End of self-monitoring intervention arm - HbA1c. - Questionnaires - Collect participant diary	+24 weeks	2 weeks after Visit 6

* Each study visit can be scheduled with +/- 2 weeks of the planned visit date.

**This data download will be for self-monitoring of blood glucose/pump data.

3 STATISTICAL PRINCIPLES

3.1 Confidence intervals and P-values

All statistical tests will use a 2-sided significance level of 5% (unless otherwise specified). All confidence intervals presented will be 95% and two sided. No adjustment for multiplicity is planned.

3.2 Adherence and protocol deviations

3.2.1 Adherence

During Visit 3, participant's adherence / tolerance of using the flash-CGM over the preceding 14 days will be assessed. To proceed with the study participant should have worn the blinded glucose monitoring device for at least 10 days during last 14 days of run-in period. If the participant fails to demonstrate adherence or develops any significant allergy or intolerance to the glucose sensor, the participant will be removed from the study.

The number and percentage of participants failing to demonstrate adherence and not wearing the blinded glucose monitoring device for at least 10 days will be presented in a table.

During the trial, intervention adherence data will be collected and analysed as part of the process evaluation and will also be considered in a sensitivity analysis of the primary outcome measure using Complier Adjusted Causal Estimation (see Section 5.2.3). During Visit 7, participant's adherence / tolerance at the control group of using the flash-CGM over the preceding 14 days will also be assessed. Participants' adherence will be defined as wearing the blinded glucose monitoring device for at least 10 days during last 14 days of Visit 7. The number and percentage of participants at the control group failing to demonstrate adherence and not wearing the blinded glucose monitoring device for at least 10 days will be presented in a table. The percentage of time wearing the glucose sensor alongside mean glucose sensor usage will also be recorded for participants in the flash-glucose monitoring arm.

3.2.2 Protocol deviations

The following are pre-defined categories for the protocol deviations:

1. Visits outside of visit window
2. Missed/delayed blood tests
3. Missed sensor download

4. Other

Protocol deviations are classified prior to unblinding of treatment. The number (and percentage) of patients with major and minor protocol deviations will be summarised by treatment group, time point as well as being summarised by the above categories. The patients that are included in the intention to treat (ITT) analysis data set will be used as the denominator to calculate the percentages. No formal statistical testing will be undertaken.

3.3 Analysis populations

All efficacy and safety analyses will be conducted following the ITT principle where all randomised participants are analysed in their allocated treatment group whether or not they receive their randomised treatment according to the protocol.

The primary outcome analysis will be conducted based on a complete-case analysis approach unless there is more than 10% of participants who have missing HbA1c at 24 weeks (or more than a 10% difference between missing data percentages in the two arms), in which case multiple imputation will be used in order to implement a more complete ITT analysis of the substantive Analysis of Covariance (ANCOVA) model (otherwise this will be performed as a sensitivity analysis, with a complete case analysis used as the primary analysis.

4 TRIAL POPULATION

4.1 Screening data

The number of patients being screened will not be reported due to discrepancies with how screening data is being collected across the different sites. The number of patients who are assessed for eligibility will be reported.

4.2 Eligibility

The number of patients identified as meeting the eligibility criteria will be presented. The number of ineligible patients, if any, will be also reported, alongside reasons for ineligibility. Reasons for any exclusion during the pre-randomisation phase will also be reported.

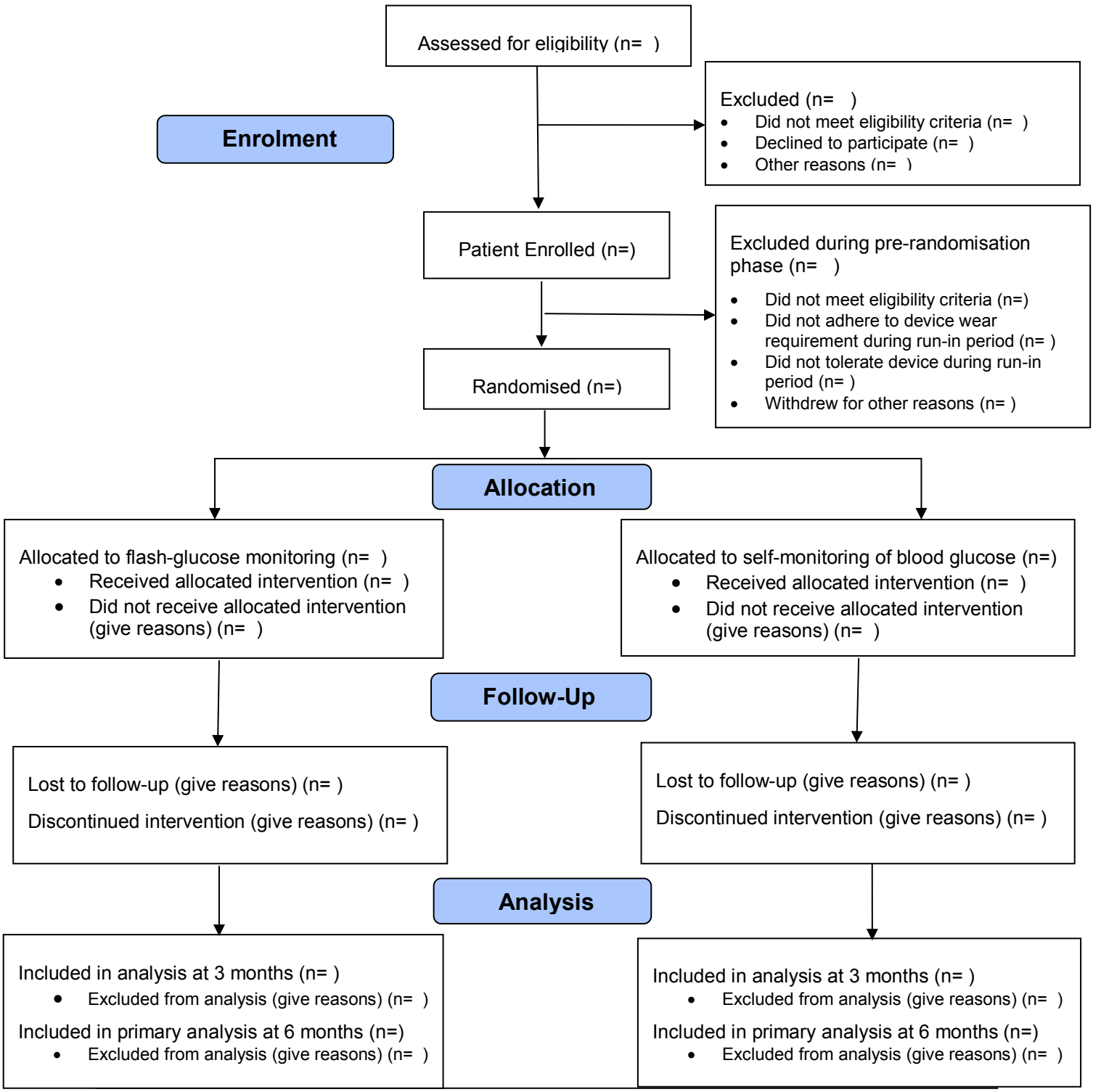
4.3 Recruitment

A CONSORT flow diagram will be used to summarise the number of patients eligible, enrolled, excluded between enrolment and randomisation, randomised, receiving their allocated

treatment, withdrawing/lost to follow-up and how many patients are included in the analysis at 3 and 6 months.

The CONSORT flow diagram template is included below in Figure 1.

Figure 1: CONSORT flow diagram template



4.4 Withdrawal/follow-up

This data will be presented in CONSORT diagram format.

4.5 Baseline patient characteristics

The following baseline characteristics will be presented descriptively, both overall and within treatment group. Continuous data will be summarised by mean and standard deviation, unless data are at least moderately skewed, in which case median and interquartile range will be used. Categorical variables will be summarised by frequencies and percentages. Statistical tests will not be conducted to test for difference between arms.

Baseline characteristics:

- Centre (Birmingham; Cambridge; Derby; Manchester; Norwich; Portsmouth; Ipswich; Wareham/Poole)
- HbA1c
- HbA1c category (7.5%-9.0%; >9.0%-11%)
- Duration of diabetes
- Treatment modality (MDI; CSII)
- Prior participation in structured education course (yes; no)
- Current use of bolus calculator (yes; no).
- BMI
- Ethnicity (White; Mixed; Asian/Asian British; Black/African/Caribbean /Black British; Other; Prefer not to answer; Not assessed)
- Deprivation decile
- Age
- Sex (male; female)
- Education (< Bachelor's degree; >= Bachelor's degree)
- Occupation (Office based; Manual; No Occupation)
- History of Diabetes complications:
    - Retinopathy (past; current; no)
    - Neuropathy: including peripheral, autonomic, and diabetic gastroparesis (yes; no)
    - Microalbuminuria or Renal impairment (yes; no)
- Use of any lipid-lowering agents (yes; no)
- Use of any anti-hypertensive agents (yes; no)
- Current Insulin therapy (Treatment modality: Multiple Daily Injection; Insulin Pump Therapy)

- History of disordered eating (yes; no)
- History of needle phobia (yes; no)
- Type 1 Diabetes Distress Scale (T1-DDS) mean-item score
- T1-DDS average-item score category (≥1.0 - 1.5; ≥1.5 - 2.0; ≥2.0 - 3.0; ≥3.0 - 6.0)[1]
- Patient Health Questionnaire (PHQ-9) total score
- PHQ-9 items sum score category (<5; ≥5 - 10; ≥10 - 15; ≥15 - 20; ≥20 – 27)[2]
- Diabetes fear of injecting and self-testing questionnaire (D-FISQ)[3] total score
- The revised Diabetes Eating Problem Survey (DEPS-R)[4] total score
- Diabetes Treatment Satisfaction Questionnaire (DTSQ)[5] total score
- The Glucose Monitoring Satisfaction Survey (GMSS)[6] mean-item score, and subscales mean-item score:
    - Openness
    - Emotional burden
    - Behavioural burden
    - Trust
- Hypoglycaemia burden assessed using:
    - Clarke score
    - Gold score
    - Number with 1 or more severe hypoglycaemia episodes in last 6 months and 12 months
- Concomitant Diabetes Medications:
    - GLP-1 analogues (yes; no)
    - Metformin (past; current)
    - SGLT2 inhibitor (yes; no)

5 ANALYSIS

5.1 Outcome definitions

5.1.1 Primary outcome

The primary outcome (endpoint) is HbA1c at 24 weeks.

5.1.2 Secondary outcomes

HbA1c based

1. HbA1c at 12 weeks
2. HbA1c ≤ 53 mmol/mol (7.0%)
    - at 12 weeks [yes/no]
    - at 24 weeks [yes/no]
3. HbA1c ≤ 59 mmol/mol (7.5%)
    - at 12 weeks [yes/no]
    - at 24 weeks [yes/no]
4. Reduction in HbA1c ≥5.5 mmol/mol (0.5%) from baseline (screening)
    - at 12 weeks [yes/no]
    - at 24 weeks [yes/no]
5. Reduction in HbA1c ≥ 11 mmol/mol (1.0%) from baseline (screening)
    - at 12 weeks [yes/no]
    - at 24 weeks [yes/no]

Sensor based

1. Time spent in the target glucose range between 3.9 to 10.0 mmol/l (70 to 180mg/dl).
2. Time spent below target glucose (<3.9mmol/l) (<70mg/dl)
3. Time spent above target glucose (10.0 mmol/l) (180 mg/dl)
4. Average glucose levels
5. Standard deviation glucose levels
6. Coefficient of variation glucose levels
7. The time with sensor glucose levels:
    - < 3.5 mmol/l (63 mg/dl)
    - < 3.0 mmol/l (54mg/dl)
    - < 2.8 mmol/l (50 mg/dl)
8. The time with sensor glucose levels in the significant hyperglycaemia (glucose levels > 16.7 mmol/l) (300mg/dl)
9. AUC of glucose below 3.0mmol/l (54mg/dl)
10. Glucose management Indicator

All the sensor-based metrics will also be analysed separately for daytime (7:00-23:00 hours) and night-time (23:00-7:00 hours) in addition to the overall period.

Non-sensor based secondary clinical

1. Daily average total insulin dose
2. Daily average basal insulin dose
3. Daily average bolus dose
4. Average number of boluses of rapid acting insulin per day

Non-sensor based secondary patient-reported (psychosocial)

1. T1-DDS: mean-item score
2. EQ-5D-5L (This outcome will not be covered in the SAP but instead will be included in the Health Economics Analysis Plan.)
3. PHQ-9 total score
4. D-FISQ total score
5. DEPS-R total score

Harms outcomes

1. Frequency of severe hypoglycaemic episodes as defined by American Diabetes Association. For purposes of analysis, a severe hypoglycaemic event will be defined as an event requiring assistance of another person actively to administer carbohydrate, glucagon, or other resuscitative actions. These episodes may be associated with sufficient neuroglycopaenia to induce seizure or coma. If plasma glucose measurements are not available during such an event, neurological recovery attributable to the restoration of plasma glucose to normal is considered sufficient evidence that the event was induced by a low plasma glucose concentration.
2. Frequency of significant ketosis events (plasma ketones >3mmol/l)
3. Nature and severity of other adverse events including the following categories:
    - *Hospital admissions*
    - *Skin reactions*

5.1.3 Process evaluation (utility and acceptability) outcomes

1. FLS2 device utilization data (FLS2 trial arm only), including:
    a. average number of scans per day (7:00-23:00 hours);
    b. average number of scans per night (23:00-7:00 hours);
    c. average number of scans over the full 24-hour period;
    d. average number of days of usage per week.

2. Average number of finger-stick glucose level tests per calendar day (as collected at visits 1, 4, 5, and 7: average over the last 14 days).
3. DTSQ total score
4. GMSS (and subscales) mean-item score

5.2 Analysis methods

All 12-week and 24-week outcome data will be presented descriptively, both overall and within treatment group, using mean (SD), median (IQR) or frequency (percentage), as appropriate. Mean (SD) change of HbA1c will also be presented overall and by treatment group. All statistical tests will use a 2-sided significance level of 5% (unless otherwise specified), and all confidence intervals will be presented at a level of 95% and will be two sided.

5.2.1 Primary outcome analysis

The primary outcome analysis will evaluate between group differences in HbA1c levels at the end of the 24-week treatment period. An ANCOVA model will be used, with 24-week HbA1c as the outcome and trial arm effect as the focus, with adjustment for baseline HbA1c and the other baseline variables included in the minimisation allocation algorithm (study centre, treatment modality, prior participation in structured education course and current use of bolus calculator) as covariates.

5.2.2 Secondary outcome analysis

HbA1c based

For the HbA1c 12-week outcome, an ANCOVA model will be used, with trial arm effect as the focus, with adjustment for baseline HbA1c and the other baseline variables included in the minimisation allocation algorithm as covariates.

For the HbA1c-based [yes/no] variables, logistic regression models will be used. For each model, trial arm effect will be the focus, with adjustment for the baseline HbA1c and the other variables included in the minimisation allocation algorithm as covariates.

Sensor-based

The respective sensor-based measures obtained during the last 2 weeks of the 24-week randomised interventions contrasting the flash-glucose against the SMBG will be compared

using ANCOVA. Analysis will be adjusted for baseline sensor values obtained during blinded run-in period, the baseline HbA1c and the other variables included in the minimisation allocation algorithm. Analysis will also be repeated for day and night-time period (the interval from 7.00 to 23:00 defines day-time period; 23:00 to 07:00 am defines the night-time period).

Non-sensor based clinical

For the insulin dose data, ANCOVA will again be used. Analysis will be adjusted for the baseline value of the outcome measure, the baseline HbA1c and the other variables included in the minimisation allocation algorithm. This analysis will be repeated to compare the two intervention groups for the insulin dose data obtained both at 12 weeks and at 24 weeks following randomisation.

Safety data, including number of severe hypoglycaemia events and number of ketone-positive hyperglycaemia, will be tabulated for all participants, including drop-outs and withdrawals, irrespective of whether CGM data are available and irrespective of whether closed-loop was operational. Severe hypoglycaemic events and ketone-positive hyperglycaemia will be tabulated in each treatment group.

Non-sensor based patient-reported (psychosocial)

ANCOVA will also be used for the psychosocial outcome data evaluation, with trial arm effect as the focus, adjusting for baseline level of the outcome and the other baseline variables included in the minimisation allocation algorithm as covariates.

In addition, analysis for caseness will be performed for the PHQ-9 and the T1DDS outcomes, by the use of logistic regression models, with trial arm effect as the focus, adjusting for baseline caseness of the outcome and the other baseline variables included in the minimisation allocation algorithm as covariates. Caseness for the PHQ-9 is defined as score of 10 or above, and for the T1DDS as 2 or above.

5.2.3 Sensitivity analyses

Sensitivity analyses will examine robustness of the primary outcome results to assumptions regarding missing data by performing multiple imputation (see Section 5.3).

The primary outcome analysis will also be repeated having included additional covariates for the consideration of improved precision. These covariates will include age, sex, BMI duration of diabetes, deprivation index quintile.

Complier-Adjusted Causal Effect (CACE) will be used, implemented using an instrumental variables approach, to estimate the causal effect of the FSL device glucose monitoring relative to finger-stick glucose monitoring.[7] Appropriate definitions of non-compliance (adherence) will be investigated within this sensitivity analysis, including the simple definition in the finger-stick monitoring arm of non-compliance equating to use of FSL device glucose monitoring for at least 10 days between the 3-month and 6-month outcome assessment.

Additional sensitivity analyses will be conducted to assess the impact of primary outcome data collected outside the visit window as specified in the protocol, that is:

- +/- 2 weeks of the planned visit date or
- +/- 4 weeks of the planned visit date.

Further sensitivity analyses will be performed to assess the potential impact of COVID-19 on the primary outcome. These will include the exclusion of participants whose treatment or outcome assessment was deemed severely impacted by COVID-19.

Finally, additional sensitivity analyses will be introduced to examine for the impact of protocol deviations or other unexpected events that may not have been captured in the protocol, including the impact on key secondary outcomes, where deemed necessary.

5.2.4 Subgroup analyses

Planned exploratory subgroup analyses will be performed for the primary outcome measure only and will include:

- Baseline HbA1c: 7.5%-9.0%; >9.0%-11%
- Treatment modality: MDI; CSII
- Prior participation in structured education course: yes; no
- Age group at recruitment (Visit 1): 16 <30; 30 to <45; 45 to <60; >=60 years
- Education: < Bachelor's degree; >= Bachelor's degree
- Hypoglycaemia Unawareness: Yes (Clarke score >3); No (Clarke score ≤3)
- Deprivation Index Quintile: (as a categorical variable, although exploratory analysis will include consideration of a linear trend or combination of adjacent categories)
- Sex: Male; Female

- Ethnic group: Categorisation will be determined by the number in individual ethnic groups. It is expected that this will take the form of white vs. non-white due to the low expected numbers in separate non-white categories, but further exploratory analyses may be performed (e.g. white vs. South Asian vs. 'other')
- PHQ-9: Mild or none depression (items sum score <10); Moderate or severe depression (items sum score ≥10).

5.3 Missing data

Should we have more than 10% missing HbA1c at 24 weeks (or more than a 10% difference between missing data percentages in the two arms), multiple imputation will be used in order to implement a more complete ITT analysis of the substantive ANCOVA model. The imputation model will include baseline and 12-week HbA1c, all the baseline variables used in the outcome model and any other recorded variables found to be predictive of missing the 24-week outcome in exploratory analyses (via a logistic regression model, with terms included using a 10% significance level). A minimum of 50 imputations will be performed, although this will be increased if the percentage of missing data is greater than expected. We will perform multiple imputation by chained equations (MICE)[8].

Complete cases analyses will be implemented for all secondary outcomes. Should we have up to 30% missing items for any subscale (or overall questionnaire if no subscales), the missing item will be imputed by the mean response of the rest items completed within participant for that particular subscale. If more than 30% missing items the response to this subscale will be set to missing, as will be the overall response of the questionnaire.

5.4 Additional analyses

The process evaluation (utility and acceptability) will include:

- Summary statistics (mean, standard deviation, range) of (i) the average number of scans per day (7:00-23:00), (ii) per night (23:00-7:00), (iii) over the full 24-hour period, and (iv) the average number of days of usage per week for the FLS2 trial arm
- Summary statistics (mean, standard deviation, and range, where appropriate) and ANCOVA will be used to explore and compare each of:
    - the average number of finger-stick glucose level tests per calendar day;
    - patient satisfaction with treatment (DTSQ total score);

- patient satisfaction with monitoring (GMSS total score; Openness subscale score; Emotional Burden subscale score; Behavioral Burden subscale score ; Trust subscale score).

by trial arm, adjusting for the baseline value of the respective outcome and the other baseline variables included in the minimisation allocation algorithm as covariates.

Additional analysis will include summary statistics (number and frequency) of initialisation of concomitant diabetes medications during the follow up period.

The analysis of the primarily *qualitative* 'Expectations and experience' questionnaire is not covered in this document. The quantitative results will be integrated into the broader process evaluation which is described more fully in the trial protocol.

5.5 Harms

All relevant adverse events (AEs) related to:

1. Frequency of severe hypoglycaemic episodes as defined by American Diabetes Association
2. Frequency of significant ketosis events (plasma ketones >3mmol/l)
3. Nature and severity of other adverse events including the following categories:
    a. Hospital admissions
    b. Skin reactions

will be reported. The number of events and the number of participants with at least 1 event across treatment arms will be reported. We will also report all individual SAEs/AEs by treatment arms. Summary measures will be frequencies and percentages (%) across treatment arms.

5.6 Statistical software

Analyses will be undertaken in Stata version 14.0[9] or later.

6 REFERENCES

1. Fisher L, Polonsky WH, Hessler DM, et al. Understanding the sources of diabetes distress in adults with type 1 diabetes. *J Diabetes Complications* 2015;29(4):572-7. doi: 10.1016/j.jdiacomp.2015.01.012 [published Online First: 2015/03/15]

2. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16(9):606-13. doi: 10.1046/j.1525-1497.2001.016009606.x [published Online First: 2001/09/15]

3. Snoek FJ, Mollema ED, Heine RJ, et al. Development and validation of the diabetes fear of injecting and self-testing questionnaire (D-FISQ): first findings. *Diabet Med* 1997;14(10):871-6. doi: 10.1002/(SICI)1096-9136(199710)14:10<871::AID-DIA457>3.0.CO;2-Y [published Online First: 1997/11/26]

4. Markowitz JT, Antisdel JE, Butler DA, et al. Brief Screening Tool for Disordered Eating in Diabetes. *Diabetes Care* 2010;33:495-500.

5. C B. The Diabetes Treatment Satisfaction Questionnaire: DTSQ. In Bradley C (Ed) *Handbook of Psychology and Diabetes: a guide to psychological measurement in diabetes research and practice. Abingdon: Routledge ISBN 9783718655625, formely published by Harwood Academic Publishers ISBN 3-7186-5562-4* 1994:111-32.

6. Polonsky WH, Fisher L, Hessler D, et al. Development of a New Measure for Assessing Insulin Delivery Device Satisfaction in Patients with Type 1 and Type 2 Diabetes. *Diabetes Technol Ther* 2015;17(11):773-9. doi: 10.1089/dia.2015.0140 [published Online First: 2015/07/15]

7. Freedman DA. Statistical models for causation: what inferential leverage do they provide? *Eval Rev* 2006;30(6):691-713. doi: 10.1177/0193841X06293771 [published Online First: 2006/11/10]

8. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine* 2011;30(4):377-99. doi: 10.1002/sim.4067

9. StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.