Statistical Analysis Plan I8R-MC-IGBM

A Simulation Study Comparing Successful Administration, Time to Administer, and User Experience of Ready-to-Use Nasal Glucagon with Reconstitutable Injectable Glucagon

NCT03765502

Approval Date: 19-Nov-2018

STATISTICAL ANALYSIS PLAN

A Simulation Study Comparing Successful Administration, Time to Administer, and User Experience of Ready-to-Use Nasal Glucagon with Reconstitutable Injectable Glucagon.

Statistical Analysis Plan Status: Final
Statistical Analysis Plan Date: 14-November-2018

Study Drug: LY900018

Sponsor Reference: I8R-MC-IGBM
Covance CRU Study: 1000071-8394076

Clinical Phase I

Approval Date: 19-Nov-2018 GMT

1. TABLE OF CONTENTS

2. ABBREVIATIONS

Abbreviations pertain to the Statistical Analysis Plan (SAP) only (not the tables, figures and listings [TFLs]).

AE	Adverse event
CRF	Case Report Form
CRU	Clinical Research Unit
CSR	Clinical Study Report
EC	Early Clinical
e.g.	For example (Latin: *exempli gratia*)
GEE	Generalized estimating equations
GEK	Glucagon emergency kit
ICH	International Conference on Harmonisation
MedDRA	Medical Dictionary for Regulatory Activities
NA	Not applicable
NASA	National Aeronautics and Space Administration
NG	Nasal glucagon
PWD	Person with diabetes
SAP	Statistical Analysis Plan
SOP	Standard Operating Procedure
SUSAR	Suspected unexpected serious adverse reaction
TFLs	Tables, Figures, and Listings
TLX	Task load index
WHO	World Health Organization

3. INTRODUCTION

This SAP has been developed after review of the Clinical Study Protocol (final version dated 21 September 2018).

This SAP describes the planned analysis of the efficacy and safety data from this study. A detailed description of the planned TFLs to be presented in the clinical study report (CSR) is provided in the accompanying TFL shell document.

The intent of this document is to provide guidance for the statistical analysis of data. In general, the analyses are based on information from the protocol, unless they have been modified by agreement between Eli Lilly and Company and Covance Early Clinical (EC) Biometrics. A limited amount of information concerning this study (e.g., objectives, study design) is given to help the reader's interpretation. This SAP must be signed off prior to first subject simulation for this study. When the SAP and TFL shells are agreed upon and finalized, they will serve as the template for this study's CSR.

This SAP supersedes the statistical considerations identified in the protocol; where considerations are substantially different, they will be so identified. If additional analyses are required to supplement the planned analyses described in this SAP, they may be performed and will be identified in the CSR. Any substantial deviations from this SAP will be agreed upon between Eli Lilly and Company and Covance EC Biometrics and identified in the CSR. Any minor deviations from the TFLs may not be documented in the CSR.

This SAP is written with consideration of the recommendations outlined in the International Conference on Harmonisation (ICH) E9 Guideline entitled Guidance for Industry: Statistical Principles for Clinical Trials[1] and the ICH E3 Guideline entitled Guidance for Industry: Structure and Content of Clinical Study Reports[2].

4. STUDY OBJECTIVES

4.1 Primary Objective

- To test the hypothesis that more Trained Users can successfully administer nasal glucagon (NG) than glucagon emergency kit (GEK) when treating a simulated episode of severe hypoglycemia.

4.2 Secondary Objectives

- To test the hypothesis that more Untrained Users can successfully administer NG than GEK when treating a simulated episode of severe hypoglycemia.
- To compare the average time to successfully administer NG, compared with GEK, for trained and untrained users.
- To compare percentages of users who complete critical administration steps for both NG and GEK, for trained and untrained users.
- To assess the number of simulations that result in successful administration of NG.
- To assess whether, overall and for trained users, persons with diabetes (PWDs) and untrained users, one device (NG or GEK) is preferred over the other.
- To assess overall, and for trained users, PWDs and untrained users, ease of use of both devices (NG and GEK).

4.3 Exploratory Objectives

- To assess participants' overall confidence in the use of NG and GEK devices.
- To compare the time it takes a PWD to teach the use of NG and GEK devices.
- To compare participants' experience of Mental Demand, Effort, Frustration and other elements of the National Aeronautics and Space Administration (NASA) Task Load Index (TLX), following NG and GEK simulations.

5. STUDY DESIGN

This is a single-center, randomized, crossover study assessing the percentage of simulations with dose successfully administered, and other experiences of administering both NG and a commercially available GEK during simulations of severe hypoglycemia emergencies. Participants in the study will administer nasal and injectable glucagon devices to manikins.

Study I8R-MC-IGBM (IGBM) is divided into 2 parts, with different types of participants: Part A (Trained Users, who will attend all study visits with a PWD) and Part B (Untrained Users). Both parts of the study may be run at the same time.

The study is intended to simulate real-life conditions to capture and document, in a controlled setting, any differences in how Trained Users and Untrained Users use NG and GEK, and their experience of doing so.

5.1 Types of Participants

Participants are defined as follows:

- **Trained User**: A participant who, during the study, will be given a level of training intended to mimic that of a real-life caregiver, relative, or close friend of a PWD.
    - "Trained User" participants must be either related to, or a close friend of, a PWD, and therefore already familiar with diabetes and the risk of severe hypoglycemia. Each Trained User must be recruited alongside the PWD with whom they are familiar, as a "dyad" (a pair of people who have a pre-existing relationship).
    - The Trained User will receive instruction on the use of each device, from the PWD, before taking part in the simulation using that device.
    - Trained Users must not have actually administered injectable glucagon (or another rescue medication) in the past. However, they may own a glucagon rescue kit, and may have been instructed in its use, as long as such instruction occurred at least 2 years before screening.
- **PWD:** The PWD who is the close friend or relative of the Trained User.
    - During the study, the PWD will be trained by site staff on the use of each device, and at least 7 days before each simulation, will relay these instructions to the Trained User.
    - The PWD previously may have been instructed in the use of a glucagon rescue kit, as long as such instruction occurred at least 2 years before screening.
- **Untrained User:** A participant who will not be given training during the study on either device, and who has limited knowledge of severe hypoglycemia and how to treat it.
    - These participants are intended to approximate a real-life colleague or acquaintance of a PWD: someone who is aware of the PWD's condition and is willing to intervene in an episode of severe hypoglycemia, but has not received instructions. Untrained Users in this study may or may not know a PWD in real life. In any case, they must not have caregiving responsibilities toward a PWD.
    - Untrained Users will participate in the study alone (not as part of a dyad).
    - Prior to the simulations in this study, Untrained Users will be given basic information about severe hypoglycemia and briefly shown an example of each rescue device.

Figure 1 (Part A) and Figure 2 (Part B) illustrate the study design.

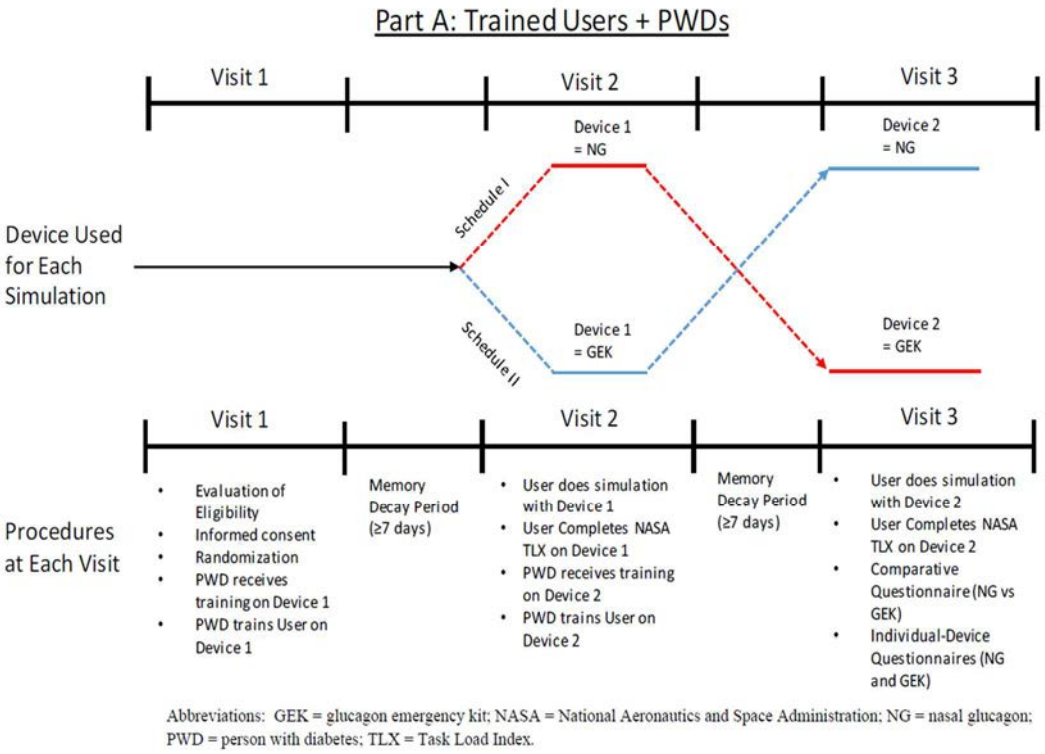

Figure 1 - Illustration of study design for Part A.

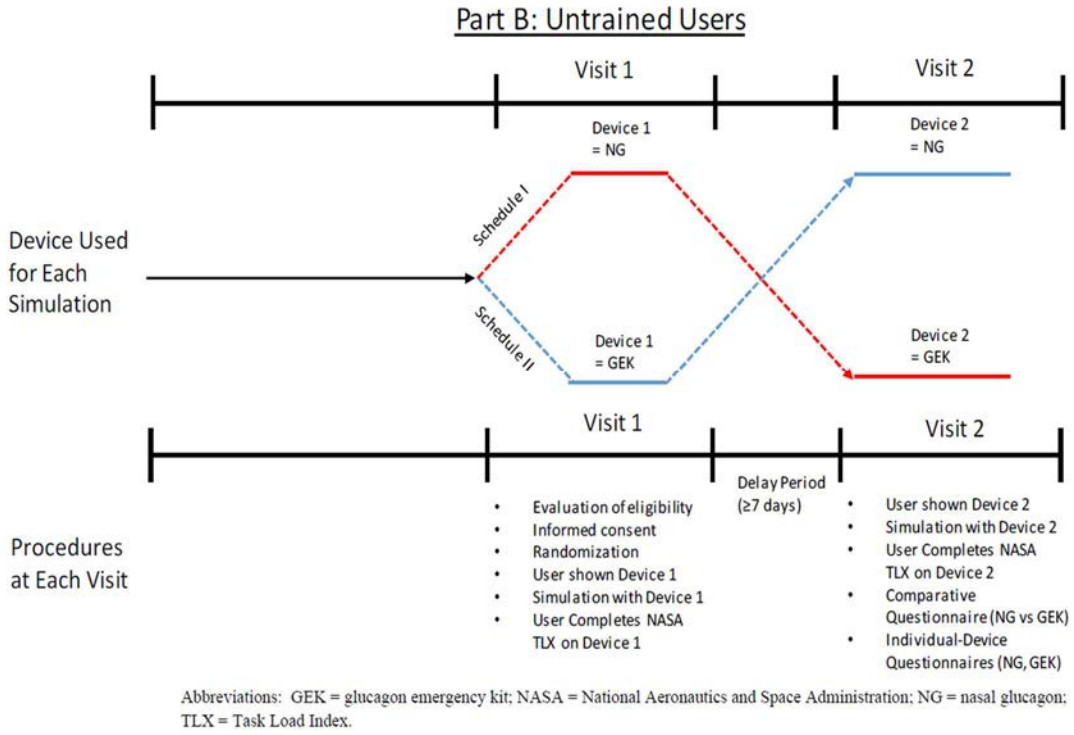

Figure 2 - Illustration of study design for Part B.

6. TREATMENTS

The following is a list of the study treatment abbreviations that will be used in the TFLs.

Part A (Trained Users who have an accompanying person with T1DM or T2DM)

Study Treatment Name	Abbreviation	Treatment order in TFL
Nasal glucagon	NG	1
Glucagon emergency kit	GEK	2

Part B (Untrained Users)

Study Treatment Name	Abbreviation	Treatment order in TFL
Nasal glucagon	NG	3
Glucagon emergency kit	GEK	4

7. SAMPLE SIZE JUSTIFICATION

The sample size is 26 in each of the 2 cohorts (Trained Users and Untrained Users). It is assumed that the percentages of successful dosing of NG and GEK will be 90% and 50%, respectively.

Additionally, a proportion of discordant pairs of 45% is assumed, which reflects a moderate within-subject correlation between the 2 treatments.

These numbers are based on the Yale 2017 study[3], where 15 out of 16 (93.8%) NG simulations were completely successful and 8/16 (50%) GEK administrations were partially or completely successful in the "caregivers" group (the equivalent of the Trained Users group in this study). Note that for GEK, the "partial or complete success" rate is used. In Yale 2017, only 2 out of 16 (12.5%) simulations were completely successful for the GEK treatment. Hence, we are using a best-case scenario approach for the successful administration rates of GEK in the current study. In conclusion, the 90% success rate for NG is based on the Yale 2017 NG success rate (15/16, 93.8%, rounded down) and the 50% success rate for GEK is based on the best-case scenario from the Yale 2017 study, i.e. partial administration of GEK in 8/16 (50%) simulations.

Based on these numbers, the sample size will provide 90% statistical power to detect a statistically significant difference between success rates of NG and GEK, using the test with a 2-sided alpha level of 0.05.

Enrollment will be continued per group until 26 Trained Users have been recruited (stratified into 13 who have an accompanying PWD with T1DM, and 13 who have a PWD with T2DM) and 26 Untrained Users have been recruited.

Participants who are randomized but do not participate in the study simulations may be replaced to ensure that enough participants may complete the study.

8. DEFINITION OF ANALYSIS POPULATIONS

The "Full Analysis Set" will consist of all subjects who participate in at least one training of a study device.

The "Per Protocol Population" will consist of all subjects who complete the study without major protocol deviations.

All protocol deviations that occur during the study will be considered for their severity/impact and will be taken into consideration when subjects are assigned to analysis populations.

9. STATISTICAL METHODOLOGY

9.1 General

Data listings will be provided for all data that is databased. Summary statistics and statistical analysis will only be presented for data where detailed in this SAP. For continuous data, summary statistics will include the arithmetic mean, arithmetic standard deviation, median, min, max, P25, P75 and N will also be presented. For categorical data, frequency count and percentages will be presented. Data listings will be provided for all subjects up to the point of withdrawal, with any subjects excluded from the relevant population highlighted. Summary statistics and statistical analyses will generally only be performed for subjects included in the relevant analysis population. For the calculation of summary statistics and statistical analysis, unrounded data will be used.

Data analysis will be performed using SAS® Version 9.4 or greater.

9.2 Demographics and Subject Disposition

Subject disposition will be listed. The demographic variables age, sex, race, ethnicity, type 1 diabetes mellitus and type 2 diabetes mellitus will be summarized and listed. All other demographic variables will be listed only. The demographic variables will also be compared, making the following comparisons:

- NG (Trained User) versus NG (Untrained User) – two sample t-test
- GEK (Trained User) versus GEK (Untrained User) – two sample t-test

9.3 Adverse Events

Where changes in severity are recorded in the Case Report Form (CRF), each separate severity of the adverse event (AE) will be reported in the listings, only the most severe will be used in the summary tables. A pre-existing condition is defined as an AE that starts before the subject has provided written informed consent and is ongoing at consent. A non-treatment emergent AE is defined as an AE which starts after informed consent but prior to dosing. A treatment-emergent AE is defined as an AE which occurs postdose or which is present prior to dosing and becomes more severe postdose.

All AEs will be listed. Treatment-emergent AEs will be summarized by treatment, severity and relationship to the study drug. The frequency (the number of AEs, the number of subjects experiencing an AE and the percentage of subjects experiencing an AE) of treatment-emergent AEs will be summarized by treatment, Medical Dictionary for Regulatory Activities (MedDRA) version 21.0 system organ class and preferred term. The summary and frequency AE tables will be presented for all causalities and those considered related to the study drug. Any serious AEs will be listed.

Associated person AE's will be listed and summarized by treatment, severity and relationship to the study drug.

9.4 Questionnaires

Trained Users and Untrained Users will complete NASA TLX questionnaires soon after each simulation, and all participants (Trained Users, Untrained Users and PWDs) will complete further questionnaires at the end of the study.

For the end-of-study questionnaires, the questionnaire directly comparing both devices will always be given first. The order of the NG- and GEK-specific questionnaires will be randomized.

NASA Task Load Index

Following each simulation, Trained Users and Untrained Users (but not PWDs) will be asked to complete Hart and Staveland's NASA TLX to gauge their experience during that simulation. The NASA TLX is a series of 21-point scales that measure various aspects of carrying out a task. Example questions include Mental Demand ("How mentally demanding was the task?"), Effort

("How hard did you have to work to accomplish your level of performance?") and Frustration ("How insecure, discouraged, irritated, stressed, and annoyed were you?").

Following each simulation, Trained Users and Untrained Users (but not PWDs) will be asked to complete Hart and Staveland's NASA TLX to gauge their experience during that simulation. The NASA TLX measures the perceived impact of carrying out a task in 6 dimensions (scales):

- Mental Demand ("How mentally demanding was the task?")
- Physical Demand ("How physically demanding was the task?")
- Temporal Demand ("How hurried or rushed was the pace of the task?")
- Performance ("How successful were you in accomplishing what you were asked to do?")
- Effort ("How hard did you have to work to accomplish your level of performance?")
- Frustration ("How insecure, discouraged, irritated, stressed, and annoyed were you?").

For each of these scales, the subject assigns as **score** using a visual analogue scale divided into 20 segments using 21 hashes. This 21-point scale (0-20) will be converted to a 0-100 point scale for the analysis.

Additionally, **weights** will be calculated for each the scales. This will be carried out using comparison cards, on which each scale is compared to the 5 other scales, creating a total of 15 comparison cards. The subject needs to choose, for each comparison, which scale was the most demanding for the task. For each time a scale was selected by the subject, a value of 1 is added to its weight. As each scale is compared to the other 5 scales exactly ones, the maximum of the weight is 5 (this scale was always selected as being the most demanding) and the minimum is zero (this scale was never selected as being the most demanding).

For the analyses, a weighted average of the 100-point-scale scores will be computed in the analysis datasets for each subject and administration method (NG / GEK).

Questionnaire Comparing Both Devices

For Trained Users and Untrained Users, the comparative questionnaire will include questions on preference, ease of use and confidence using each device.

Person with diabetes will receive a differently worded questionnaire that will include questions on their preference of device, and their impression of the user experience for ease of use and confidence.

Device-Specific Questionnaires (NG and GEK)

All participants (Trained Users, Untrained Users and PWDs) will answer 2 device-specific questionnaires, regarding NG and GEK individually. These will include questions related to ease of use and confidence in using each device, without comparing to the other device.

All questionnaire data will be listed and summarized by part and treatment. For Part A (Trained Users), summaries will also be stratified between those users who have an accompanying PWD with T1DM, those who have an accompanying PWD with T2DM and overall.

9.5 Statistical Methodology

The following endpoints will be summarized by treatment using descriptive statistics and listed:

- The percentage of Trained Users and Untrained Users that perform a successful administration
- The time taken by Trained Users and Untrained Users to successfully administer NG or GEK
- The percentages of Trained Users and Untrained Users who complete all of the critical steps
- The percentage of all simulations that result in a successful administration of NG, for all users combined
- The average length of time taken by PWDs to teach the use of NG and GEK to their associated Trained User,

The McNemar test will be used to assess the percentage of Trained Users that perform a successful administration. The exact-sign test will be used as a sensitivity analysis.

Successful administration will be assessed, for each cohort:

- NG (Trained User) versus GEK (Trained User)
- NG (Untrained User) versus GEK (Untrained User)

Example SAS code for the McNemar test

```
proc freq data = success;
   weight count / zeros;
   tables NG_trained*GEK_trained / agree;
   exact mcnem;
run;
```

Example SAS code for the exact-sign test

```
proc univariate data=success;
  var success_diff;
run;
```

Where `success_diff` is 1 if NG is successful and GEK is not, 0 if both treatments are successful or not successful and -1 if NG is not successful and GEK is successful.

For Part A (Trained Users), assessment of successful administration will also be stratified between those users who have an accompanying PWD with T1DM, and those who have an accompanying PWD with T2DM.

Time to administer a dose will be assessed using summary statistics, making the following comparisons:

- NG (Trained User) versus NG (Untrained User) – two sample t-test
- GEK (Trained User) versus GEK (Untrained User) – two sample t-test

- NG (Trained User) versus GEK (Trained User) – Paired t-test
- NG (Untrained User) versus GEK (Untrained User) – Paired t-test

This analysis, using descriptive statistics and the appropriate t-test, will be assessed for successful and unsuccessful administration combined, as well as successful administration only. Time needed to complete glucagon administration will be shown categorically based on the following categories:

- <30 seconds
- 30 seconds-45 seconds
- 45 seconds-1 minute
- 1 minute-1 minute 15 seconds
- 1 minute 15 seconds-1 minute 30 seconds
- 1 minute 30 seconds-1 minute 45 seconds
- 1 minute 45 seconds-2 minutes
- 2 minutes-2 minutes 15 seconds
- 2 minute 15 seconds-2 minutes 30 seconds
- 2minutes 30 seconds-2 minutes 45 seconds
- 2 minutes 45 seconds-3 minutes
- 3-4 minutes
- 4-5 minutes
- 5-10 minutes
- >10 minutes

A graphical presentation of the successful administration times will be provided using a Kaplan-Meier plot.

To simultaneously model the training effect (Trained Users versus Untrained Users) and the treatment effect (NG versus GEK), a generalized estimating equations (GEEs) model will be performed, with treatment, training and the treatment-by-training interaction effect as independent variables, and the participant as participant identifier. As each Trained User and Untrained User will perform both the NG and the GEK administration, the data are paired for the device type but not paired for the level of training. Using a GEE model will enable estimation of both effects simultaneously, and also estimation of the interaction effect, which will test whether the difference between devices varies between Trained Users and Untrained Users. Furthermore, the randomization order (NG-GEK versus GEK-NG) can be tested within this model. Given the low sample size of this study and multitude of variables that can be entered, this GEE modelling will be exploratory in nature.

Primarily, a GEE model using the binary distribution (success versus no success) and the identity link will be produced: this will produce estimated rate differences with 95% confidence intervals. In case this model fails to converge, log or logit links can be attempted, to estimate relative risks or odds ratios, respectively.

Example SAS code:

```
proc genmod data=xx descending;
class  subjid arm trt ftrt ;
model success = arm|trt ftrt / link=identity dist=bin;
repeated subject=subjid;
lsmeans  arm*trt / ilink  diff;
lsmeans  arm / ilink  diff;
lsmeans  trt / ilink  diff;
lsmeans  ftrt / ilink  diff;
run;
```

where arm = trained or untrained, trt = NG or IM and ftrt = first treatment received

9.6 Other Assessments

All other assessments not detailed in previous sections will be listed but not summarized or statistically analyzed.

10. INTERIM ANALYSES

No interim statistical analyses are planned.

11. CHANGES FROM THE PROTOCOL SPECIFIED STATISTICAL ANALYSES

There were no changes from the protocol specified statistical analyses.

12. REFERENCES

1. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, ICH Harmonized Tripartite Guideline, Statistical Principles for Clinical Trials (E9), 5 February 1998.

2. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, ICH Harmonized Tripartite Guideline, Structure and Content of Clinical Study Reports (E3), 30 November 1995.

3. Yale JF, Dulude H, Egeth M, Pichè CA, Lafontaine M, Carballo D, Margolies R, Dissinger E, Shames AR, Kaplowitz, N, Zhang MX, Zhang S, Guzman CB. Faster use and fewer failures with needle-free nasal glucagon versus injectable glucagon in severe hypoglycemia rescue: a simulation study. Diabetes Technol Ther. 2017;19(7):423-432.

13. DATA PRESENTATION

13.1 Derived Parameters

N and percentage values should be reported as whole numbers. Median values should be treated as an observed parameter and reported to the same number of decimal places as minimum and maximum values.

13.2 Missing Data

Missing data will not be displayed in listings.

13.3 Insufficient Data for Presentation

Some of the TFLs may not have sufficient numbers of subjects or data for presentation. If this occurs, the blank TFL shell will be presented with a message printed in the center of the table, such as, “No serious adverse events occurred for this study.”