Imperial College London

Self-Assessment Method for Statin side effects Or Nocebo trial (SAMSON)

Statistical Analysis Plan
Version 1.0

NCT02668016

January 14, 2020

Prepared by:

Statistics Collaborative, Inc.

CONFIDENTIAL

Imperial College London

SAMSON
Statistical Analysis Plan

January 14, 2020

Prepared by: Statistics Collaborative, Inc.

Author: _____ Date: 14-Jan-2020

Katrina Epnere, MPH
Biostatistician II

_____ Date: 14-Jan-2020

Janet Wittes, PhD
President

Approved by: Imperial College London

_____ Date: _____

Darrel Francis, M.D.

CONFIDENTIAL

Contents

Exhibit

Abbreviations

AE	Adverse event
ATC	Anatomical Therapeutic Chemical
CI	confidence interval
CRF	case report form
EQ-5D-3L	EuroQol
ICH GCP	International Conference on Harmonisation Good Clinical Practice
ICTU	Imperial Clinical Trials Unit
MedDRA	Medical Dictionary for Regulatory Activities
PT	preferred term
SAMSON	self-assessment method for statin side effects or nocebo trial
SAP	Statistical analysis plan
SCI	Statistics Collaborative, Inc.
SD	standard deviation
SOC	system organ class
TSQM	Treatment Satisfaction Questionnaire for Medicine
WHO	World Health Organization
WHO-DD	World Health Organization Drug Dictionary

1. Introduction

This statistical analysis plan (SAP), which is based on protocol Version 1.2 dated August 1, 2019, defines the methods and analyses that Imperial College London plans to use to analyze the data from the Self-Assessment Method for Statin side effects Or Nocebo trial (henceforth, SAMSON). This study adheres to the principles outlined in the Medicines for Human Use (Clinical Trials) Regulations 2004 (SI 2004/1031), amended regulations (SI 2006/1928), and the International Conference on Harmonisation Good Clinical Practice (ICH GCP) guidelines. It is being conducted in compliance with the protocol, the Data Protection Act, and other regulatory requirements as appropriate. If the protocol is subsequently amended, this SAP may be amended as well. Should the SAP and the protocol be inconsistent with respect to the planned analyses, the language of the SAP is governing.

2. Investigational plan

2.1. Study design

SAMSON is a randomized controlled trial evaluating in a cohort of participants the proportion of adverse symptoms truly due to the effect of statin rather than nocebo. The investigators plan to randomize participants to receive, in a random predefined order, the study interventions. Each participant will receive 12 sets of pre-labelled HDPE containers. Four sets of containers will contain no medication; four will contain a one-month supply of matched placebo; and four will contain a one-month supply of atorvastatin 20mg. At the start of the next calendar month after the screening visit, the participants will commence the trial intervention. The research nurse will call participants to remind them to start on the first day of the next month after screening; thereafter, the participants will receive a monthly reminder to switch to the next set of HDPE containers. Each day participants will rate their daily symptom on a phone application; they will complete three additional questionnaires each month.

2.2. Study objectives and outcome measures

Front-line clinicians cannot currently test whether symptoms an individual participant experiences are the pharmacological result of a statin or due to other phenomena (e.g., nocebo). The value of such a tool would be two fold:

- The tool would allow individual participants to establish for themselves whether they truly suffered a side effect from the drug or are victims of nocebo – which may in fact be more common.
- By separating the components, the tool would permit clinical researchers to explore the determinants of each, opening opportunities to obtain better clinical outcomes.

The main objectives of this study are to the following:

- To develop a method for determining for an individual participant what extent experienced symptoms are associated with the statin or merely reflect a nocebo effect.
- In a cohort of participants who have stopped statins because of adverse symptoms, to evaluate in what proportion the symptoms are truly due to the statin.

2.3. Study hypotheses

The study aims to test the following hypotheses:

- Hypothesis 1: More than 30% of participants enrolling into the study will complete it.
- **Hypothesis 2:** Overall, more than 50% of symptom burden is nocebo rather than pharmacological.

  The effect of nocebo effect on side effects will be defined as:

  $$Nocebo\ effect = \frac{Nacebo\ component}{Total\ side\ effect\ (Pharmacological+Psychological)} = \frac{(average\ symptoms\ score\ on\ placebo - average\ symptom\ score\ on\ no\ medication)}{(average\ symptom\ score\ on\ statin - average\ symptom\ score\ on\ no\ medication)}$$

- **Hypothesis 3:** At six months after completion of the study, at least 50% of participants will either be taking statins or have declined statins for reasons other than perceived side effects.

2.4. Randomization

Eligible participants are enrolled on InForm, which allocates each participant a random predefined order during which to take the study interventions. The Imperial Clinical Trials Unit (ICTU) generates these random codes and supplies them to the production pharmacy.

3. Study schedule

3.1. Study days

Potential participants attend a screening visit during which the study doctor receives written informed consent from those who decide to participate. The study doctor then assesses the participants' eligibility for the study by evaluating their past medical history and previous statin intolerance. Participants' blood pressure is measured. Participants who do not have a recent lipid profile recorded in the last 6-months are offered the option of having one undertaken as part of the screening visit. The study doctor determines if the participant is suitable to be enrolled in the study. Suitable participants are enrolled on InForm.

Scheduled follow-up telephone calls are undertaken each month during the 12-month period of the trial. The study nurse monitors the participants' scores. If the scores show severe discomfort or if participants are not scoring on their phone, the study nurse makes unscheduled telephone follow-up calls and, if required, arranges an unscheduled study visit to see the study doctor who performs unscheduled tests deemed necessary.

The end of study will be defined to occur when the specified number of participants have been recruited, all participants have completed the 18-month phone interview, and the database is locked.

The 12-month follow-up contact may be combined with the end of study visit; if so this would be a face-to-face visit at the study center. If so, the month 12 visit must be completed before the scheduled unblinding. The end of study visit may take place up to 31 days after the 12-month telephone follow-up. The Exhibit presents the study visit schedule

Exhibit. Visit schedule

	D1	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M18
	Screening and Enrolment	Telephone follow-up	Telephone follow-up	Telephone follow-up	Telephone follow-up	Telephone follow-up	Telephone follow-up	Telephone follow-up	Telephone follow-up	Telephone follow-up	Telephone follow-up	Telephone follow-up	End of Study Visit	6-month Follow-up visit
Informed consent	X													
Inclusion/exclusion	X													
Demography	X													
Medical history	X													
Blood pressure	X													
Lipid profile (optional)	X													
Interview	X	X	X	X	X	X	X	X	X	X	X	X	X	X
EuroQol	X	X	X	X	X	X	X	X	X	X	X	X	X	
TSQM	X	X	X	X	X	X	X	X	X	X	X	X	X	
Confounding life events questionnaire	X	X	X	X	X	X	X	X	X	X	X	X	X	

3.2. Unscheduled visits

Unscheduled assessments are not performed unless participants develop adverse events that the chief investigator considers 'related' to the trial procedure or to atorvastatin therapy.

4. Sample size determination and power calculation

4.1. Sample size calculation – Hypothesis 2

Each "effect of nocebo" will be a value which, for the sake of the calculation in the protocol, was assumed to be between 0 and 1. The study aims to report an average effect of nocebo for

the population that has a 95% CI of ±10%. To achieve this, assuming a conservative scenario of individual-participant values scattered uniformly from 0 to 1 (i.e., SD = $1/\sqrt{12} = 0.29$), the number of participants to be studied will need to be $\geq N$ where $0.29/\sqrt{N} \leq 0.10/1.96$(i.e., $N > (1.96 \times 0.29/0.10)^2 = 36$. The investigators planned to recruit 50 participants. In all a total of 62 participants were screened and 60 were randomized.

5. Statistical analysis: general considerations and conventions

Descriptive and inferential statistics will be used to summarize results of the SAMSON study. Continuous variables will be summarized using the number of subjects (N), mean, SD, median, 25th and 75th percentiles, and minimum and maximum. Discrete variables will be summarized using counts and percentages.

Summaries will be provided for demographics, medical history, concomitant medications, and adverse events. For summaries of medical history, concomitant medications, and AEs, the Medical Dictionary for Regulatory Activities (MedDRA®) and the World Health Organization Drug dictionaries, as appropriate, will be used.

All data listings, summaries, and statistical analyses will be generated using SAS® Version 9.4 (or higher) or other validated software.

6. Efficacy analysis

As described in Section 2.3, the study has three hypotheses.

6.1. Primary outcome (Hypothesis 2)

Hypothesis 2: Overall >50% of symptom burden is nocebo rather than pharmacological.

For the trial, all participants receive a smartphone or, if they prefer, the application can be downloaded to their existing phone. Access to the phone will allow them to record real-time daily documentation of symptoms experienced on a visual analogue scale of 0-100. Participants will rate symptoms every day, with the daily scores averaged into a monthly score.

Each participant receives four months (4 × 28 days) each of no medication, placebo, and active medication (i.e, statin). We represent the individual participant's symptom score data as follows:

$$\begin{matrix} x_{0_1} & x_{p_1} & x_{s_1} \\ x_{0_2} & x_{p_2} & x_{s_2} \\ x_{0_3} & x_{p_3} & x_{s_3} \\ x_{0_4} & x_{p_4} & x_{s_4} \end{matrix}$$

$$\bar{\boldsymbol{x}}_{\boldsymbol{0}} = \frac{\sum_{j=1}^{4} x_{0\,j}}{4} = \frac{\sum_{j=1}^{4}\left(\left(\sum_{i=1}^{28} x_{0_i}\right)/28\right)}{4}, \bar{\boldsymbol{x}}_{\boldsymbol{p}} = \frac{\sum_{j=1}^{4} x_{p\,j}}{4} = \frac{\sum_{j=1}^{4}\left(\left(\sum_{i=1}^{28} x_{p_i}\right)/28\right)}{4}, \bar{\boldsymbol{x}}_{\boldsymbol{s}} = \frac{\sum_{j=1}^{4} x_{s\,j}}{4} = \frac{\sum_{j=1}^{4}\left(\left(\sum_{i=1}^{28} x_{s_i}\right)/28\right)}{4}$$

$\bar{x}$- mean symptom score for no medication, placebo, statin, $\boldsymbol{i}$-day 1 to 28 of treatment month, $\boldsymbol{j}$ – treatment month 1 to 4

We assume no carry-over effect. Thus, the estimated mean of the effect of nocebo is calculated as follows:

$$Mean\ nocebo\ effect = \sum_{n=1}^{N}\left(\frac{\bar{x}_{p_n} - \bar{x}_{0_n}}{\bar{x}_{s_n} - \bar{x}_{0_n}}\right)/N$$

N= *number of study participants*

Because the study participants are independent of each other, the variance of the estimated mean can be calculated as the as the sum of variances of the individual participants:

$$Variance\ of\ the\ mean\ nocebo\ effect = \boldsymbol{s^2} = \sum_{n=1}^{N} {s_n}^2 = \sum_{n=1}^{N} Var\left(\frac{\bar{x}_{p_n} - \bar{x}_{0_n}}{\bar{x}_{s_n} - \bar{x}_{0_n}}\right)$$

To calculate the variance of the estimate, we use a Taylor series approximation, which is a function of the variance of the numerator, variance of the denominator, and the covariance of the two:

1) $Var(numerator) = Var\left(\bar{x}_p - \bar{x}_0\right) = {s_p}^2 + {s_0}^2 - 2r_{s_p,s_0}s_p s_0$

2) $Var(denominator) = Var\left(\bar{x}_s - \bar{x}_0\right) = {s_s}^2 + {s_0}^2 - 2r_{s_s,s_0}s_s s_0$

3) $Cov(numerator,\ denominator) = cov(\bar{x}_p - \bar{x}_0\ ,\ \bar{x}_s - \bar{x}_0) = cov\ (\bar{x}_p,\ \bar{x}_s) - cov\ (\bar{x}_p,\ \bar{x}_0) - cov\ (\bar{x}_0,\ \bar{x}_s) + cov\ (\bar{x}_0,\ \bar{x}_0) = cov\ (\bar{x}_p,\ \bar{x}_s) - cov\ (\bar{x}_p,\ \bar{x}_0) - cov\ (\bar{x}_0,\ \bar{x}_s) + {s_0}^2$

$$where, for\ example, \qquad cov(\bar{x}_p, \bar{x}_s) = \frac{\sum_{j=1}^{4}(x_{p_j} - \bar{x}_p)(x_{s_j} - \bar{x}_s)}{4-1}$$

4) $${s_n}^2 = Var\left(\frac{\bar{x}_p - \bar{x}_0}{\bar{x}_s - \bar{x}_0}\right) = \left(\frac{\bar{x}_p - \bar{x}_0}{\bar{x}_s - \bar{x}_0}\right)^2 \left(\frac{Var\ (\bar{x}_p - \bar{x}_0)}{(\bar{x}_p - \bar{x}_0)^2} + \frac{Var\ (\bar{x}_s - \bar{x}_0)}{(\bar{x}_s - \bar{x}_0)^2} - 2\frac{cov(\bar{x}_p - \bar{x}_0\ ,\bar{x}_s - \bar{x}_0)}{(\bar{x}_p - \bar{x}_0)\ (\bar{x}_s - \bar{x}_0)}\right)$$

s *= standard deviation,* $\boldsymbol{s^2}$*= variance, r = correlation*

The Taylor's series approximation is based on asymptotic theory. Should the data not allow stable estimates of the components of the variance, we will use a bootstrapped estimate of the variance.

To calculate the 95% CI for the mean nocebo effect, the test statistic is $t = \sqrt{N}(\mu - \mu_0)/\sqrt{s^2}$

The null hypothesis is rejected if $t > t_{\alpha,N-1}$ where, in SAS, $t_{\alpha,N-1}$ is TINV(1-α, N-1);

In SAS, the p value associated with the test statistic T is PVAL=1-probt(abs(T), N-1) or PVAL=cdf('T', T, N-1).

The 95% confidence interval is $Mean\ nocebo\ effect \pm t_{\alpha/2,N-1}\ \sqrt{s^2}\ /\sqrt{N}$

6.1.1. Missing and partial data for primary outcome

For the calculation of mean symptom score for particular treatment (statin, placebo, no medication) missing scores (i < 28) will be handled by only using study months that have at least 10 of daily symptom scores non-missing. We will perform sensitivity analyses using different thresholds for non-missing scores (such as 14 and 16 non-missing scores) and compare the percentage of missing symptoms scores during the 'pill' months (statin or placebo) versus 'no pill' months. We will assess the temporal pattern of missing symptom scores by comparing the percentage of missing scores at the beginning of the study to the percentages at later study months.

The nocebo effect for an individual participant will be calculated only if he or she has non-missing symptom scores for at least one month of each treatment (j ≥ 1). If the participant has not recorded any symptom score for one of the three treatments (j=0 for statin, placebo, or no pill) the nocebo effect cannot be calculated, and the primary endpoint analysis will exclude that participant.

6.2. Other efficacy outcomes

6.2.1. Hypothesis 1

Hypothesis 1: More than 30% of enrolled participants will complete the study.

1) Point estimate $\hat{p} = x/N$ where x is the number of participants who completed the study and N is the number of enrolled participants
2) Hypotheses: $H_0 : \hat{p} = p_0 \ \ and \ \ H_A = \hat{p} > p_0 \qquad where\ p_0 = 0.3$
3) Test statistic $z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/N}}$
4) Reject null hypothesis if $z > z_\alpha, where\ z_{0.05} = 1.645$
5) 95% confidence interval:$\hat{p} \pm z_{a/2}\sqrt{\hat{p}(1-\hat{p})/N} = \hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/N}$

6.2.2. Hypothesis 3

Hypothesis 3: At six months after completion of the study, at least 50% of the participants will either be taking statins or have declined statins for reasons other than perceived side effects. This hypothesis will be tested in the same way as Hypothesis 1 above but setting $p_0 = 0.5$.

7. Characteristics of the population

7.1. Demographics and baseline characteristics

Quantitative variables will be summarized using the number of subjects (N), mean, SD, median, 25th and 75th percentiles, and minimum and maximum. Qualitative variables will be summarized using counts and percentages.

The following demographic and baseline characteristics will be summarized: age, gender, height, weight, ethnicity, and past use of statin. Blood pressure and selected biochemistry laboratory measurements reported as well.

Age in years will be calculated as the integer portion of the following:

[(Date of randomization- Date of birth) + 1] / 365.25.

Unless otherwise stated, percentages will be calculated relative to the number of subjects randomized.

7.2. Concomitant medications

Information on concomitant medication may be summarized by generic name as reported on the concomitant medication CRF. The number and percentage of subjects who took at least one drug within each generic type will be presented. Subjects will be counted only once if they take the same generic medication more than once.

7.3. Medical history

Medical history will be coded using the MedDRA and summarized by system organ class (SOC) and preferred term (PT). Medical history may be sorted by descending overall frequency, by SOC and PT, in the summary tables.

8. Protocol deviations

Important protocol deviations will be summarized by the category reported on the CRF form and sorted by descending overall frequency

9. Adverse events

Adverse events (AEs) are monitored throughout the study and documented on the appropriate AE form. They are coded using the MedDRA dictionary by system organ class (SOC) and preferred term (PT); they will be classified by seriousness, severity, and relation to study medications.

AEs will be summarized by subject, not event. AE data will be also presented in listings.

10. Quality of life measures and confounding life events

Each month participants fill out the following two validated questionnaires related to the impact of their side-effects on their quality of life: EuroQol (EQ-5D-3L) and the side effects domain of the Treatment Satisfaction Questionnaire for Medicine (TSQM) questionnaire. Participants will also fill in a short questionnaire detailing any potentially confounding life events they experienced over the previous month (e.g., change of daily routine, holidays, bereavement). At the end of study visit participants will have an exit interview exploring the nature of symptoms occurring during the study in case they may differ from those described in the baseline interview.

The daily quality of life scores will have their distribution described by the mean and standard deviation or, if not normally distributed, by the median and interquartile range. These scores will then be aggregated to monthly average scores. The correlation within group and intraclass correlation will be used to assess within-subject agreement and group mean reliability. To establish the measurement properties (i.e., convergent validity and measurement invariance) of the new measure we will examine the correlations between the month averages and each of the two monthly validated scoring systems (EQ-5D-3L and TQSM). A strong correlation ($r>0.4$) will reflect satisfactory convergent validity. The stability of the correlations across the 12 months will reflect the degree of measurement invariance. We will test this formally by a pair of path analysis models using M*plus*: one model in which the within-time correlations between the monthly aggregate of the new measure and the validated measure (i.e. EQ-5D-3L and TQSM) are fixed across time, and another where they are allowed to differ. If the latter model does not offer a significant improvement in model fit (assessed by chi-squared and fit indices), this suggests measurement invariance.