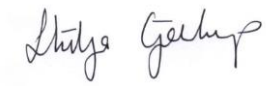
ORaCIES Trial

Statistical Analysis Plan

VERSION 1.0

10 JULY 2018

Author	
Name:	Dr Federico Ricciardi
Role:	Statistician
Signature:	
Date:	05/07/2018

Author	
Name:	Dr Hülya Gökalp
Role:	Statistician/Research Fellow Helped writing parts related to Data Analysis
Signature:	
Date:	05.07.2018

Approver	
Name:	Professor Patrick Stone
Role:	Chief Investigator
Signature:	
Date:	10-Jul-2018

1 Contents

2 Study Summary 3
3 Introduction 4
3.1 Purpose of the Statistical Analysis Plan 4
3.2 Authorship 4
3.3 Organisation of Data and Analyses 4
4 Summary of Quantitative Trial Data 5
4.1 Observation times 5
4.2 Summary of Outcome Measures 5
4.2.1 Primary outcome 5
4.2.2 Secondary outcomes 5
4.3 Summary of demographic and course details 6
5 Analyses 6
5.1 Recruitment and Retention 6
5.2 Protocol Deviations 6
5.3 Description of Baseline Variables 7
5.4 Drop-out 7
5.5 Primary Outcome Analysis 7
5.5.1 Model Checking 8
5.6 Secondary Analyses 8
6 References 9

2 Study Summary

Title: Online Randomised controlled trial to improve Clinical Estimates of Survival (ORaCIES).

Design: A web-based, multi-site, randomised controlled, trial of an online training resource to improve the recognition of dying in palliative care patients.

Aims: Primary:

To assess if an online training resource can help medical students to formulate estimates of the probability of dying for palliative care patients that are more similar to experts' estimates.

Secondary:

1) To assess if any effect of the intervention is maintained after two weeks.
2) To evaluate if the factors that medical students use to make their decision, alter after the intervention, and become more similar to the experts.
3) To assess the expertise (the ability to discriminate and be consistent in their decision) of the medical students before and after the intervention.

Outcome measures: Primary outcome:

- A continuous estimate provided from the students for each vignette; ranging from 0% (no chance of death within 72 hours) to 100% (certain death within 72 hours).

Secondary outcomes:

1. Maintenance effect, measured by using the primary outcome measure at the two week follow up time.
2. Cue weighting of the individual students as compared to the experts; that is the coefficient of each factor as part of their "judgement policy".
3. Level of expertise assessed with the Cochran-Weiss-Shanteau (CWS) score.

Population: Participants will be medical students from up to 33 accredited courses in the UK.

Eligibility: Participants must be 18 or more, enrolled on a UK medical course and must be in the penultimate or final year of the course. Participants must be able to read English.

3 Introduction

3.1 Purpose of the Statistical Analysis Plan

This document contains details of the main quantitative, statistical, analyses for the "ORaCIES" trial. These analyses shall be pre-specified in order that they are not influenced by the collected trial data after unblinding. The statistical analysis plan (SAP) does not preclude the undertaking of further, ad-hoc, analyses, although the results of any such further analyses should be interpreted carefully. Furthermore, the SAP does not prevent the adaptation of any part of the trial analysis, should situations arise in which such adaptation is deemed necessary. Any such adaptation shall be fully justified and transparent.

The SAP contains details of quantitative analyses only and does not describe any qualitative and/or economic analyses.

3.2 Authorship

The SAP has been written by Dr. Federico Ricciardi, with the collaboration of Dr. Hülya Gökalp.

3.3 Organisation of Data and Analyses

Unblinding of collected data shall occur after completing the statistical analysis. The programs and code to be used for statistical analyses shall be prepared, where possible, prior to the unblinding of data. Two statisticians shall perform the analysis relating to the primary outcome independently, in order to ensure its accuracy.

Prior to performing analyses, basic checks shall be performed on the collected, blinded, data to ensure accuracy. Each outcome (primary and secondary) variable and baseline demographic variable shall be checked for:

- missing values;
- values beyond an acceptable range;
- other inconsistencies.

If missing values or other inconsistencies are present the corresponding data shall be checked with the aid of the researchers and, where necessary, either corrected or deemed to be missing. Any such changes made to the unblinded dataset shall be documented fully.

4 Summary of Quantitative Trial Data

4.1 Observation times

The times at which data are collected during the trial are as follows:

a) Baseline (t=0, immediately after consent);
b) After randomisation (t=1);
c) After intervention (t=2);
d) Follow up (t=3, 2 weeks after intervention).

Outcomes for different sets of vignettes are recorded at every time point but the first one. Table 1 provides a summary of the times at which measures are collected. The data recorded at time points a), b) and c) above will constitute the dataset for the purpose of analysis of the primary outcome.

4.2 Summary of Outcome Measures

4.2.1 Primary outcome

The primary outcome will be the continuous estimate provided from the students for each vignette; ranging from 0% (no chance of death within 72 hours) to 100% (certain death within 72 hours).
A sample of 128 subjects (64 subjects in each group) is required to detect a medium effect size (Cohen's d = 0.5) between the intervention and control groups, assuming a common standard deviation, 80% power and using a two sample t-test at the 5% significance level. We estimate that it will be necessary to recruit approximately 183 subjects in order to obtain a final sample size of 128 participants with complete data sets for analysis. The anticipated 30% drop-out rate (i.e. participants who start but do not complete the task) has been estimated on the basis of previous similar studies by our own group.

4.2.2 Secondary outcomes

1. Maintenance effect measured by using the primary outcome measure at the two week follow up time point.
2. Cue weighting of the individual students as compared to the experts; that is the coefficient of each factor as part of their "judgement policy".
3. Level of expertise: The level of expertise will be assessed with the Cochran-Weiss-Shanteau (CWS) score [1-2]. This score details the level of ability to discriminate and the level of consistency. This will help us to understand if the participants become better at discriminating after the intervention, and if these decisions become more consistent.

4.3 Summary of demographic and course details

Furthermore, a collection of basic medical and personal information shall be made at baseline. This information shall include:

- Age and Gender;
- Ethnicity;
- Place and year of study.
- Palliative care experience (training, confidence, placements, experience)

These variables will be analysed descriptively.

Table 1 Outcome measures and time points at which outcomes are.

Outcome measure	Baseline (t=0)	After Randomisation (t=1)	After Intervention (t=2)	Follow-Up (t=3)
Palliative Care Experience	X			
Age + Gender	X			
Ethnicity	X			
Place of Study	X			
Year of Study	X			
Probabilities of dying for the first set of vignettes (30 + 10 repeated cases)		X		
Probabilities of dying for the second set of vignettes (20 + 6 repeated cases)			X	
Probabilities of dying for the second set of vignettes (20 + 6 repeated cases)				X

5 Analyses

5.1 Recruitment and Retention

A CONSORT diagram shall be presented to provide a detailed description of participant numbers at each time point during the trial.

5.2 Protocol Deviations

Protocol deviations that could impact on the results of the analyses will be described (for example, details concerning participants withdrawing consent to continue with the study). All individuals who withdraw consent shall be excluded from the analyses from the point of

withdrawal, although any data collected from such participants prior to the point of withdrawal shall be included unless otherwise specified by the participant.

5.3 Description of Baseline Variables

Baseline measures for the two treatment orders will be summarised by treatment assignment and overall, giving means and standard deviations for Age and frequencies and percentages for categorical and binary variables.

5.4 Drop-out

Based on past experience, 30% drop-out rate has been anticipated. Reasons for withdrawal (if known) and time point at which withdrawal took place will be reported.

5.5 Primary Outcome Analysis

The primary outcome, measured on a continuous scale, will be the estimate provided from the students for each vignette; ranging from 0% (no chance of death) to 100% (certain death).

Let the i subscript denote the i^{th} student, the j subscript denote the j^{th} vignette, with $j \in (1, \dots, J^{(t)})$ where $t \in \{1, 2, 3\}$ is the time point of interest and $J^{(t)}$ the total number of vignettes for time point t

- $X_{ij}^{(t)}$ = estimated probability of dying for the j^{th} vignette, by the i^{th} student at time t;
- $T_j^{(t)}$ = experts' estimate for the j^{th} vignette at time t.

For each student, to estimate the degree of agreement between the novices' estimates of the probability of dying and the experts' reference values at each time point, we will calculate the Mean Absolute Difference [3] as:

$$M_i^{(t)} = \frac{\sum_{j=1}^{J^{(t)}} |X_{ij}^{(t)} - T_j^{(t)}|}{J^{(t)}}.$$

A linear regression model for the agreement at time 2 will be fitted. The model shall include an adjustment for baseline mean absolute difference and treatment group. Hence, the linear model for the primary outcome is given by:

$$M_i^{(2)} = \beta_0 + \beta_1 Z_i + \beta_2 M_i^{(1)} + \varepsilon_i,$$

where

- Z_i is the treatment assigned to the i^{th} student;
- ε_i is the error term.

A p-value pertaining to the hypothesis test of

$$H_0: \beta_1 \geq 0 \text{ versus } H_1: \beta_1 < 0$$

shall be reported. A P-value ≥ 0.01 shall be reported to two decimal places; a P-value in the range 0.001-0.01 shall be reported to three decimal places and a P-value less than 0.001 shall be reported as 'P-value < 0.001'.

In addition to this, we will also calculate the degree of agreement using the Bland Altman method [4]. This will be done pre and post the intervention.

5.5.1 Model Checking

The model for the primary outcome analysis includes an assumption that the primary outcomes, $M_i^{(2)}$, are normally distributed. The normality of the primary outcome variable shall be assessed through the construction of appropriate histograms and normal quantile-quantile plots. If such plots suggest that the primary outcome variable is not normally distributed, then appropriate transformations of the primary outcome variable shall be considered.
The normality of the estimated error terms shall be assessed using normal quantile-quantile plots. The homoscedasticity of the same estimated error terms shall be assessed using a scatter plot of the residuals. Possible influential observations and outliers shall be identified. Sensitivity to such influential observations and/or outliers (if present) shall be considered.

5.6 Secondary Analyses

The maintenance of the study effect will be studied similarly to the primary outcome analysis with the Mean Absolute Differences from the two week follow up time point (i.e., $M_i^{(3)}$) as dependent variable and adding $M_i^{(2)}$ as independent variable to control for. The linear model for the analysis will hence be:

$$M_i^{(3)} = \alpha_0 + \alpha_1 Z_i + \alpha_2 M_i^{(1)} + \alpha_3 M_i^{(2)} + \varepsilon_i,$$

To assess the individual "judgement policy" and its eventual improvement, a linear model for each one of the students will be fitted, using the estimated probability of dying as dependent variable and the values of the different cues as independent covariates. The Experts' and student's coefficients shall be compared in a descriptive fashion.

For each participant, the CWS score, and the subcomponents (e.g. consistency and discrimination) will be calculated at each time point by using the repeated vignettes. The CWS score is calculated as the ratio of discrimination and inconsistency. The discrimination refers to the individual participant's differential assessment of the various vignettes in a set, and is calculated as the variance among responses to different vignettes. The inconsistency refers to the individual participant's assessment of the same vignette similarly over time, and is calculated as the variance among responses to the same vignette. The higher the CWS score, the more consistent and discriminating the judge is.

6 References

1. Weiss DJ, Shanteau J. CWS: A user's guide. Retrieved; 2005.

2. Weiss DJ, Shanteau J., Harries, P. People Who Judge People. J. Behav. Dec. Making, 19 (2006), pp. 441–454

3. Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim. Res., 30 (2005), pp. 79-82.

4. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. The lancet. 1986;327(8476):307-10.