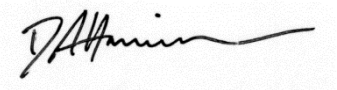
Provision Of Psychological support to People in Intensive care

Psychological Outcomes following a nurse-led Preventative Psychological Intervention for critically ill patients (POPPI) trial

Statistical Analysis Plan

Version 1.1, 28/11/2017

REC number:	15/SC/0287
Trial sponsor:	ICNARC
Trial sponsor reference:	ICNARC/01/03/14
Trial funder:	NIHR HS&DR Programme
Funder reference:	12/64/124
ISRCTN number:	ISRCTN53448131
NIHR CRN Portfolio ID number:	18940
Author:	Dr Jerome Wulff

Role, Name and Position	Signature	Date
Chief investigator: Prof Kathryn Rowan (Director of Scientific & Strategic Development/CTU Director, ICNARC)		28/11/2017
Senior Statistician: Dr David Harrison (Head Statistician, ICNARC)		28/11/2017

Version history

Version number	Date	Summary of main changes from previous versions
1.0	10/08/2017	N/A
1.1	27/11/2017	Minor typographical corrections and changes to references; inclusion of baseline/resource use covariates in MI (4.1.4); addition of adherence variable to MI model (4.1.4)

Abbreviations

AIC	Akaike Information Criteria
BIC	Bayesian Information Criteria
CAM-ICU	Confusion Assessment Method for the Intensive Care Unit
CEA	Cost Effectiveness Analysis
CI	Confidence Interval
Cluster-RCT	Cluster-Randomised Controlled Trial
CMP	Case Mix Programme
EQ-5D	EuroQol 5-dimension quality of life questionnaire
GCS	Glasgow Coma Scale
GLMM	Generalised Linear Mixed Model
HADS	Hospital Anxiety and Depression Scale
HRQoL	Health Related Quality of Life
ICNARC	Intensive Care National Audit and Research Centre
IMD	Index of Multiple Deprivation
IPAT	Intensive care Psychological Assessment Tool
ITT	Intent-to-treat
NHS	National Health Service
NMB	Net Monetary Benefit
POPPI	Psychological Outcomes following a nurse-led Preventative Psychological Intervention for critically ill patients
PSS-SR	PTSD Symptom Scale – Self Report version
PTSD	Post-Traumatic Stress Disorder
QALYs	Quality Adjusted Life Years
RASS	Richmond Agitation Sedation Scale
RCT	Randomised Controlled Trial
SAP	Statistical Analysis Plan
STAI	State Trait Anxiety Inventory

Table of contents

1. Background and rationale

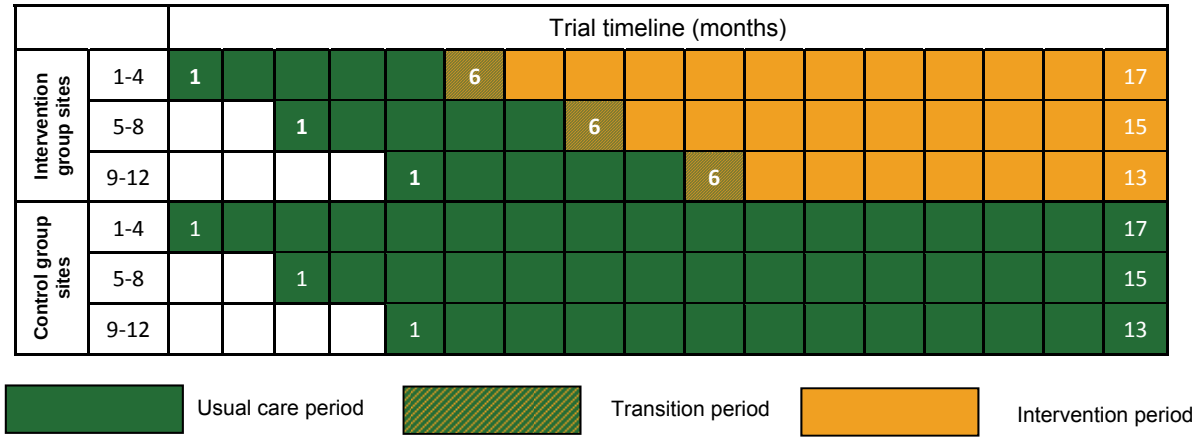
The POPPI (Psychological Outcomes following a nurse-led Preventative Psychological Intervention for critically ill patients) trial ("the Trial") is a cluster-randomised controlled trial (cluster-RCT) comparing a complex nurse-led preventative psychological intervention with usual care in reducing patient-reported post-traumatic stress disorder (PTSD) symptom severity, and other reported psychological morbidities at six months.

The study design (Figure 1) is of 24 sites, randomly assigned to either intervention or control (usual care) groups, each recruiting for between 13 and 17 months, with a staggered start to allow for roll-out of the intervention. The end of the Trial will be when the final patient has completed their six months follow-up.

The purpose of this Statistical Analysis Plan (SAP) is to document the planned analyses to be carried out to support the completion of the Final Report to the study funder and for inclusion in manuscripts for publication in the scientific literature. Additional exploratory analyses, not necessarily identified in this SAP, may also be performed. Any post-hoc or unplanned analyses not identified in this SAP will be clearly identified as such in the respective Report/manuscript.

This SAP has been agreed in advance of inspecting any outcome data from the intervention period of the Trial, so that data-derived decisions in the analyses are avoided.

Figure 1. Cluster-RCT schedule



2. Aim and objectives

2.1. Aim

The aim of the Trial is to evaluate the clinical and cost-effectiveness of a complex nurse-led preventative psychological intervention in reducing patient-reported PTSD[1] symptom severity, and other reported psychological morbidities at six months.

2.2. Objectives

The specific objectives are:

i. To evaluate the effect of the complex intervention on patient-reported PTSD symptom severity and other psychological morbidities and quality of life at six months; and
ii. To estimate, in an integrated economic analysis, the cost-effectiveness of the intervention.

An integrated process evaluation will be conducted to assess the fidelity and quality of the implementation of the intervention, and identify important contextual factors to better understand how the intervention works.

3. Methods

3.1. Trial design

Parallel group cluster-RCT, with staggered opening and a baseline (pre-intervention) period.

3.2. Setting

Twenty-four NHS adult, general critical care units in the UK ("sites").

3.3. Inclusion and exclusion criteria

The inclusion and exclusion criteria of this study are as described below.

3.3.1. Eligibility criteria for sites (clusters)

The following criteria must be met for a site to participate in the Trial. A site must:

i. show that recruitment to target, timely data collection, and delivery of the complex intervention are feasible - via completion of a site feasibility questionnaire;
ii. commit to dedicate adequate resources to carry out the complex intervention;
iii. agree to adhere to randomisation into either the control group or the intervention group;
iv. have two Joint Principal Investigators (PIs) identified to lead POPPI at the site (a lead nurse and a lead clinician);
v. agree, where possible, to recruit all eligible patients to POPPI and to maintain a POPPI Screening Log to include reasons why eligible patients were not recruited
vi. agree to use the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)[2] for assessing delirium and Richmond Agitation Sedation Scale (RASS)[3] for assessing sedation status for the duration of the trial; and
vii. be actively participating in the Case Mix Programme (CMP) – the national clinical audit for critical care units coordinated by ICNARC.

Sites that have taken part as an intervention site in the POPPI Feasibility Study (ISRCTN61088114) were not be eligible for selection.

3.3.1. Inclusion criteria for patients

Patients must meet all of the following criteria:

viii. age 18 years or greater;
ix. greater than 48 hours in the critical care unit;
x. receipt of Level 3 critical care (for any period of time) during first 48 hours in the critical care unit;

xi. between +1 and -1 on the RASS;
xii. Glasgow Coma Scale score of 15;
xiii. English-speaking; and
xiv. ability to communicate orally.

3.3.2. Exclusion criteria for patients

Patients must not meet any of the following criteria:

i. pre-existing chronic cognitive impairment, such as dementia;
ii. pre-existing psychotic illness, such as schizophrenia;
iii. pre-existing chronic posttraumatic stress disorder;
iv. receiving end-of-life care; or
v. previously recruited to POPPI.

3.4. Outcomes

All outcomes will be assessed and reported at the individual patient level.

3.4.1. Primary outcomes

The primary outcome for the clinical evaluation will be patient-reported PTSD symptom severity at six months, measured using the PTSD Symptom Scale – Self Report version (PSS-SR), which conforms to all DSM-IV diagnostic criteria for PTSD and which has been validated for use in critical care unit survivors.[4]

The primary outcomes for the economic evaluation will be incremental costs, quality-adjusted life years (QALYs) and net monetary benefit at six months.

3.4.2. Secondary outcomes

The secondary outcomes will be:

i. days alive and free from sedation to day 30;
ii. duration of critical care unit stay;
iii. PSS-SR greater than 18 points at six months;
iv. depression at six months, measured using the Hospital Anxiety and Depression Scale (HADS)[5];
v. anxiety at six months, measured using the HADS[5]; and
vi. health-related quality of life (HRQoL) at six months, measured by the EuroQol (EQ-5D-5L) questionnaire.

3.5. Power calculation

The initial power calculation for the POPPI cluster-RCT was calculated for the original grant submission and prior to conducting the POPPI feasibility study. It was based on very limited data to inform it – available at that time – namely, routine non-specific (with respect to the proposed POPPI trial population) data from the ICNARC Case Mix Programme (the national clinical audit for adult critical care in the UK) and more specific outcome data but only from a single-centre study of 100 patients. Despite this, to ensure a smooth transition from the POPPI feasibility study to the POPPI cluster-RCT (in the eventuality that feasibility was demonstrated), the initial pre-feasibility study power calculation formed the basis for the original ethics application for the POPPI cluster-RCT.

Following completion of the POPPI feasibility study, the assumptions underlying the initial pre-feasibility study power calculation were reviewed using the results from the feasibility study to ensure the proposed design retained adequate power – to produce the pre-cluster-RCT power calculation. The amount of additional information on which to update the assumptions, however, remained small – with only two critical care units having participated in the RCT processes and procedures feasibility study (providing information on the outcome measure) and a further two critical care units having participated in the delivery of the intervention feasibility study (providing information on rates of consent and patients assessed as being at high risk).

Finally, during the early phase of recruitment to the POPPI cluster-RCT, the assumptions underlying the pre-cluster-RCT power calculation were reviewed again once outcome data became available from the baseline (pre-intervention) period for 20 (of the 24) sites.

Details of these three stages are set out below.

3.5.1. Initial pre-feasibility study power calculation

The original POPPI cluster-RCT design, prior to conducting the POPPI feasibility study, was for 24 sites each recruiting eligible admissions for eleven months. The eleven months consisted of a five-month baseline period during which both intervention and control sites delivered usual care, a one-month transition period (to be excluded from the primary analysis of the cluster-RCT) during which intervention sites were trained and began to deliver the intervention, and a five-month intervention period during which intervention sites delivered the intervention. Control sites continued to deliver usual care throughout the baseline, transition and intervention periods. This design was selected to provide at least 90% power,

based on the method of Hussey and Hughes for a general, multi-period, cluster-randomised controlled trial[6] with a type I error rate of 0.05 and based on the following assumptions:

- a mean of 14 points and standard deviation of 12 points for the PSS-SR (primary outcome measure) for control group patients and for intervention group patients during the baseline period – estimated from patients receiving usual care in a previous single-centre study[7];
- an estimated intra-cluster correlation (ICC) of 0.254 – estimated, as there was no multicentre data available for the PSS-SR, by making a conservative assumption of 0.5 for the between-site coefficient of variation[8] (corresponding to a between-site standard deviation of 7 points);
- a detectable treatment effect of a reduction of 4 points on the PSS-SR based on a difference between groups equivalent to the reliable change index for the PSS-SR[9] (of 8 points) being observed in 50% of eligible patients assessed as being at high risk of psychological morbidity using the IPAT[10] in intervention sites during the intervention period; and
- an estimated harmonic mean of the number of patients completing follow-up of 76 per site per annum (corresponding to 32 in each five-month period) – estimated using data from the ICNARC Case Mix Programme for potentially eligible patients admitted to adult, general critical care units across England, Wales and Northern Ireland, assuming 10% mortality at six months following recruitment and 80% follow-up among survivors.

It was anticipated that, with the above design and assumptions, the estimated total number of patients recruited would be 2,904 (based on Case Mix Programme data). Staged roll-out in three staggers, each of eight sites (four intervention and four control) two months apart, was planned solely for practical delivery of the training for the intervention.

The above initial pre-feasibility study power calculation was included in the original trial protocol submitted for ethical approval (submitted during the feasibility study due to the need to transition rapidly from feasibility study to cluster-RCT) and was in place at the start of recruitment to the POPPI cluster-RCT.

3.5.2. Pre-cluster-RCT power calculation

Following completion of the feasibility study and prior to the start of recruitment to the cluster-RCT, the assumptions underlying the initial pre-feasibility study power calculation were reviewed using results from the feasibility study, resulting in the following assumptions:

- a mean of 6 points and standard deviation of 7.5 points for the PSS-SR (primary outcome measure);

- an estimated ICC of 0.138 – estimated by retaining a conservative assumption of 0.5 for the between-site coefficient of variation (corresponding to a between-site standard deviation of 3 points);
- a detectable treatment effect of a reduction of 2.9 points on the PSS-SR based on a difference between groups equivalent to a re-estimated reliable change index for the PSS-SR (of 8.6 points) being observed in 40% of eligible patients assessed as being at high risk of psychological morbidity using the IPAT, with 16% of recruited patients declining the intervention;
- an estimated harmonic mean of the number of patients completing follow-up of 52 per site per annum (corresponding to 22 in each five-month period) – re-estimated using data from the ICNARC Case Mix Programme for potentially eligible patients admitted to the 24 critical care units participating in the POPPI cluster-RCT, and retaining the assumptions, supported by data from the feasibility study, of 10% mortality at six months following recruitment and 80% follow-up among survivors.

This power calculation review established that the planned design retained greater than 90% power under these revised assumptions. It was anticipated that, with the above design and assumptions, the estimated total number of patients recruited would be 1,914 (based on Case Mix Programme data) in the 24 sites.

3.5.3. Final review of assumptions in pre-cluster-RCT power calculation

During the early phase of recruitment to the cluster-RCT, the day-to-day Trial Management Group noted that the recruitment rate was below anticipated. A decision was taken, in consultation with the Independent Chairs and members of the Trial Steering Committee and the Data Monitoring and Ethics Committee, to undertake a further review of the assumptions underlying the pre-cluster-RCT power calculation once outcome data were available for patients recruited during the five-month baseline period in both intervention and control sites. This review, undertaken using data available on 9 August 2016 (in month 12 of study recruitment), identified:

- a mean of 10.3 points and standard deviation of 10.8 points for the PSS-SR (primary outcome measure);
- an ICC of 0.087 (95% confidence interval 0 to 0.192) for the PSS-SR;
  [with mean, standard deviation and ICC estimated using all available data from the previous observational study, the feasibility study and the baseline period of the cluster-RCT]
- a detectable treatment effect of a reduction of 4.2 points on the PSS-SR – estimated by retaining the same effect size as a multiple of the within-site standard deviation;

- an harmonic mean of the number of patients completing follow-up of 30.7 per site per annum (corresponding to 12.8 in each five-month period) – estimated using observed data from the baseline period.

This review of assumptions established that the planned design had an anticipated 78% power under the observed parameter estimates (and, allowing for uncertainty in the between-site variation, between 73% and 85% power).

Consequently, the decision was taken to extend recruitment in stagger 1 and 2 sites to the end of planned recruitment in stagger 3 sites (corresponding to an harmonic mean of 16.5 patients completing follow-up per site during the intervention period, allowing for the variation from five to nine months duration across staggers). With this extension to recruitment, the planned design had an anticipated 85% power (and, allowing for uncertainty in the between-site variation, between 79% and 91% power). It was anticipated that, with this extension to recruitment, the estimated total number of patients recruited would be 1,378.

Recruitment continued to be monitored closely to ensure 1,378 (or more) patients were recruited and, to ensure this, a further extension to recruitment for an additional two months in all sites was approved by the Independent Chairs and members of the Trial Steering Committee and the Data Monitoring and Ethics Committee.

3.6. Allocation of sites

Participating sites were allocated to intervention or control groups using a restricted randomisation approach. A full enumeration approach to minimising imbalance[11] was selected to ensure balance across the arms in geographical location, teaching status and size of unit. Balance on geographical location was ensured by grouping the sites within each stagger according to location. We performed simulations of alternative ways to balance on size of unit comparing:

i. Balancing on teaching status and number of beds
ii. Balancing on teaching status and number of level 3 admissions
iii. Balancing on teaching status, number of beds and number of level 3 admissions

The best combination of balancing on the above three factors (balance on teaching status and number of level 3 admissions) was used to perform the final random allocation. Each stagger was made up of 8 sites and allocated 4 each to the intervention and control groups. The site allocations were done during the second month of recruitment for each stagger, on 3 November 2015, 16 December 2015 and 17 February 2016, respectively.

4. Statistical methods

4.1. General analysis issues

4.1.1. Analysis population

All analyses will be based on the intention to treat principle. The patients will be analysed according to the group they were randomised to, irrespective of whether the treatment allocated was received.

4.1.2. Sequence of planned analyses

All final, planned, analyses identified in the protocol and in this SAP will be performed only after the last patient has completed his/her treatment and the outcome measures have been recorded. A blinded data review meeting may be held prior to database lock and completion of the final analyses. In addition, the database will not be unlocked, random code unblinded or analyses completed until this SAP has been approved.

As the duration of follow-up for the primary outcome (6 months) is long relative to the duration of recruitment (intervention period between 7 and 11 months), no interim analysis of effectiveness was planned.

4.1.3. Analysis software

Analyses will be performed using Stata/SE Version 14.2 for Windows 64-bit x86-64 (StataCorp LP, College Station, Texas, USA). Multiple imputation will be performed in R Version 3.3.2 (The R Foundation for Statistical Computing, Vienna, Austria).[12]

4.1.4. Methods for withdrawals and missing data

All the patients who provided informed consent will be accounted for in the report of the Trial. Mortality at six months is anticipated to be 10% and loss to follow-up for the primary outcome is anticipated to be 20% among survivors. Loss-to follow-up for mortality at six months is anticipated to be <1%. Patients that withdraw from the trial and do not give permission for data collected prior to withdrawal will be used in the final analysis, those that die before six months and those lost to follow-up for mortality will be excluded from the analysis of six month psychological outcomes. Patients recruited during the transition period will also be excluded from the analysis. All other recruited patients will be included in the primary analysis, with outcomes imputed.

Loss to follow-up will be reported by treatment group. Reasons for withdrawal and loss to follow-up will be reported, when known.

Multiple imputation will be used to complete missing baseline and resource use covariates and non and partial responses for the PSS-SR, HADS and EQ-5D-5L, under the assumption that responses are missing at random (MAR) conditional on the observed data.[13] Two-level imputation (patients nested in sites) will be implemented using the 'jomo' package in R.[14] The overall scores on each measure will be imputed, not individual item responses. The imputation model will include the following covariates:

- Site level covariates (* denotes covariates used to balance treatment allocation):
    - Teaching status of hospital (teaching, non-teaching)*
    - Number of beds in the critical care unit (linear)
    - Number of critical care unit admissions receiving Level 3 care staying at least 48h during the pre-trial period,1 April 2014 to 31 March 2015 (linear)*
    - Allocated treatment group (intervention, control)
- Patient level covariates:
    - Time period (baseline, intervention) and interaction with treatment group
    - Age in years (linear)
    - Gender (female, male)
    - Ethnicity (white, non-white)
    - Quintile of Index of Multiple Deprivation (IMD) 2015[15] (categorical)
    - Documented pre-existing anxiety and/or depression prior to hospital admission (anxiety, depression, both, none)
    - Planned admission to the critical care unit following elective/scheduled surgery (yes, no)
    - ICNARC Physiology Score[16] from the first 24h following admission to the critical care unit (linear)
    - Last National Early Warning Score (NEWS)[17] prior to consent (linear)
    - Health-related quality of life (HRQOL) at time of consent, assessed as health thermometer score from 0 to 100 (linear)
    - Short-form State-Trait Anxiety Inventory (STAI-6)[18] at time of consent, scored from 6 to 24 (linear)
    - Duration of stay in the critical care unit in days (linear)
    - Number of days of delirium, as assessed by the CAM-ICU[2], in the critical care unit (linear)
    - Number of days receiving sedatives/anxiolytics/anaesthetics in the critical care unit (linear)
    - Number of days receiving sleep medications in the critical care unit (linear)

- Receipt of benzodiazepines in the critical care unit (yes, no)
- Number of days receiving antipsychotics in the critical care unit (linear)
- Number of days receiving analgesics in the critical care unit (linear)
- Number of days receiving antidepressants in the critical care unit (linear)
- Number of days receiving vasoactive agents in the critical care unit (linear)
- Number of days receiving mechanical ventilation in the critical care unit (linear)
- Duration of stay in hospital following discharge from the critical care unit (linear)
- Adherence to intervention (binary)
- PSS-SR at six months (linear)
- HADS at six months (linear)
- EQ-5D-5L at six months (linear)

Twenty multiply imputed datasets will be generated using Markov Chain Monte Carlo (MCMC) drawing a sample every 1000 iterations, following an initial 1000 iteration burn-in. The random number seed will be set to 6627.

For the primary clinical and cost effectiveness outcomes two sensitivity analyses will be used to address alternative assumptions regarding the missing data mechanism: missing completely at random (MCAR) and missing not at random (MNAR).

To evaluate the results under the assumption of MCAR, the analyses will be repeated using complete case data (i.e. only those patients returning a completed questionnaire).

To evaluate the results under the assumption that responses are MNAR, i.e. the probability of missing data depends on the patient's outcome after conditioning on the observed data; a pattern-mixture model approach[19] will be used. Pattern-mixture models allow the outcome to be modelled differently according to whether it is observed or missing. To inform the assumptions about the parameters for the missing pattern that cannot be estimated from the data (sensitivity parameters), expert opinion about outcome differences between patients with missing versus complete data will be elicited from a representative sample of the clinical staff involved with the POPPI trial across the different trial centres and other interested experts.[20]

4.1.5. Data transformation

If applicable, appropriate method of transformation (e.g. log, squared, cubic, square root, etc.) will be use to transform non-normally distributed continuous variables.

4.1.6. Multiple comparisons and multiplicity

No adjustment will be made to account for multiple endpoints or multiple subgroups; P<0.05 will be taken to represent a statistically significant result. The results of subgroup analyses will be interpreted taking into account the number of significant findings that would have been expected by chance alone.

4.2. Statistical analyses

4.2.1. Screening and recruitment

Screening, recruitment and follow-up will be presented in the form of a CONSORT diagram, based on the CONSORT extension for cluster-randomised trials.[21]

Descriptive statistics will be performed using the screening logs completed by all the participating sites during patient recruitment period. Patients' data recorded which will be summarized are as follows:

1. Total patients admitted
2. Total patients who stayed >48 hrs (Yes/No) – n (% of total admitted)
3. Total patients completed screening (patients with a final status) – n
4. Total patients not completed screening (patients without final status) – n
5. Patients who met stable criteria (Yes/No) – n (% of total completed screening)
    a. Reason did not meet stable criteria:
        i. No level 3 care in 1st 48 hrs – n (% of those not meeting stable criteria)
        ii. Not aged ≥ 18 yrs – n (% of those not meeting stable criteria)
        iii. Not English speaking – n (% of those not meeting stable criteria)
        iv. Previous recruited to POPPI – n (% of those not meeting stable criteria)
        v. Pre-existing chronic cognitive impairment – n (% of those not meeting stable criteria)
        vi. Pre-existing chronic PTSD – n (% of those not meeting stable criteria)
        vii. Pre-existing psychotic illness – n (% of those not meeting stable criteria)
6. Met daily transient criteria (Yes/No) – n (% of those meeting stable criteria)
   (To work out Yes: Met stable criteria = Yes AND Final status = Not eligible)
7. Reason did not meet transient criteria:
    a. Able to communicate orally - n (% of those not meeting transient criteria

b. Between +1 and -1 on the RASS - n (% of those not meeting transient criteria)
c. GCS of 15 - n (% of those not meeting transient criteria)
d. Not receiving end of life care - n (% of those not meeting transient criteria)
e. Able to consent - n (% of those not meeting transient criteria)

8. Potentially eligible patients (enrolled, refused, eligible not enrolled, other AND eligibility unknown)
   a. Missed – n (%)
   b. Eligibility unknown (%)
9. Approached (enrolled AND refused consent) (Yes/No) – n (% of potentially eligible)
   a. Enrolled – n (% of approached)
   b. Refused consent – n (% of approached)
10. Patient level indicators (to be produced overall and per month):
   a. How many times each patient underwent daily screening
      i. When screening ended for each patient (date of admission + day last screened)
   b. Percentage of days screening not occurring (e.g. weekends)

4.2.2. Demographic and baseline characteristics

Baseline demographic and clinical data will be summarised for the ITT population, for each of the two treatment groups in each of the two time periods. Continuous variables will be summarized as mean (standard deviation) and median (interquartile range) whilst categorical variables will be summarized as number (percent). There will be no statistical testing for any of the summary measures whilst comparing the baseline variables between the treatment groups. The following baseline variables will be compared between the two treatment groups.

i. Age in years
ii. Gender (female, male)
iii. Ethnicity (white, mixed, Asian, black, other, not stated)
iv. Quintile of IMD 2015 (1=least deprived to 5=most deprived)
i. Documented pre-existing anxiety/depression (anxiety, depression, both, none)
ii. Planned admission to the critical care unit following elective/scheduled surgery (yes, no)
iii. ICNARC Physiology Score from the first 24h following admission to the critical care unit
iv. APACHE II score from the first 24h following admission to the critical care unit
v. Duration of stay in the critical care unit prior to consent

vi. Number of days experiencing delirium in the critical care unit prior to consent
vii. Last NEWS prior to consent
viii. STAI-6 at time of consent
ix. HRQOL at time of consent (health thermometer score)

4.2.3. Treatments received in the critical care unit

Treatments received in the critical care unit will be summarised for the ITT population, for each of the two treatment groups in each of the two time periods. Treatments received will be summarised as number (percent) of patients receiving the treatment, the median (interquartile range) number of days on which the treatment was received (among those receiving the treatment) and the mean (standard deviation) number of days on which the treatment was received (for all patients, including those that did not receive the treatment). There will be no statistical testing for any of the summary measures whilst comparing the treatment variables between the treatment groups. The following treatment variables will be compared between the two treatment groups:

i. Sedatives/anxiolytics/anaesthetics
ii. Sleep medications
iii. Benzodiazepines (note that benzodiazepines will also be included as either sedatives/anxiolytics/anaesthetics or sleep medications, as appropriate)
iv. Antipsychotics
v. Analgesics
vi. Antidepressants
vii. Vasoactive agents
viii. Mechanical ventilation

4.2.4. Delivery of the intervention

Uptake of the POPPI Online Training will be reported for intervention sites over time as the percentage of the enumerated critical care unit staff that had completed the training course by month against a target of >80% completion.

Delivery of the intervention at a patient level will be summarised for patients in the intervention group during the intervention period. The following will be reported for all patients:

i. Number (percent) of patients consenting to assessment using the Intensive care Psychological Assessment Tool (IPAT)
ii. Among those consenting, number (percent) of patients assessed using the IPAT
iii. Median (interquartile range) IPAT score
iv. Number (percent) of patients with IPAT score ≥ 7

The following will be reported for patients with IPAT score ≥ 7:

v. Number (percent) of patients by number of stress support sessions received (0, 1, 2, 3)
vi. Reasons for not receiving all three stress support sessions
vii. Number of patients receiving tablet computer (percent of those receiving stress support session one)
viii. Number of patients reporting using tablet computer (percent of those receiving tablet computer)
ix. Numbers of patients receiving Relax and Recover DVD and Getting well, staying well booklet (percent of patients receiving stress support session two)

4.2.5. Clinical effectiveness analysis – primary outcome

The primary analysis for the clinical evaluation will examine if there is a significant difference in the mean PSS-SR at six months between patients recruited to the intervention group compared to the control group using a generalised linear mixed model (GLMM) at the individual patient level (patients nested within sites and within treatment group/time period).

The model will include the following terms:

- Fixed effects at the site level (* denotes covariates used to balance treatment allocation):
    - Teaching status of hospital (teaching, non-teaching)*
    - Number of beds in the critical care unit (linear)
    - Number of critical care unit admissions receiving Level 3 care staying at least 48h during the pre-trial period,1 April 2014 to 31 March 2015 (linear)*
    - Allocated treatment group (intervention, control)
- Fixed effects at the patient level:
    - Time period (baseline, intervention) and interaction with treatment group
    - Age in years (restricted cubic splines, 4 knots)
    - Gender (female, male)
    - Ethnicity (white, non-white)
    - Quintile of IMD 2015 (categorical)
    - Documented pre-existing anxiety and/or depression prior to hospital admission (anxiety, depression, both, none)
    - Planned admission to the critical care unit following elective/scheduled surgery (yes, no)
    - ICNARC Physiology Score from the first 24h following admission to the critical care unit (restricted cubic splines, 4 knots)

- Random effects (intercepts) at the following levels:
    - Site

The identity link (i.e. linear regression) will be used as the link function for the model and robust variance estimation[22] will be used to estimate the standard errors of the covariates as it adjusts for possible deviations from the model's assumptions. Rubin's rules will be used to combine estimates from the multiply imputed datasets. The coefficients with their 95% confidence intervals (CI) and p-values will be presented for the fixed effect covariates whilst only the coefficients with their 95% CI will be reported for the random effect variables. The primary effect estimate will be the interaction (difference in difference) between treatment group and time period. Similar models will be developed for the secondary outcomes.

A secondary analysis will use structural mean models with an instrumental variable of randomised allocated treatment to estimate the efficacy (adherence adjusted causal effect) of the stress support sessions among those patients consenting to psychological assessment and stress support sessions, assessed as being at high risk of psychological morbidity (IPAT score ≥ 7) and receiving at least two stress support sessions.[23]

A sensitivity analysis allowing the missing PSS-SR to be MNAR will use Bayesian pattern-mixture models, consistent with the specification for the primary analysis. All priors will be 'minimally informative', except those governing the differences between the observed and missing outcomes which will be informed by expert opinion. The sensitivity of the results to a full range of diversity of opinion will be examined through a comparison of pooled and individual priors. Posterior probabilities and 95% credible intervals will be reported.

4.2.6. Clinical effectiveness analysis – secondary outcomes

Analyses of the secondary outcomes will also be performed using GLMMs (like the primary outcome analysis), with identity link (i.e. linear regression) for continuous secondary outcomes (reported as difference in means with 95% CI and p-value) and logit link (i.e. logistic regression) for binary secondary outcomes (reported as odds ratio with 95% CI and p-value). Robust variance estimation method will be used to estimate the standard errors of the covariates in both the mixed linear and logistic regression models.

4.2.7. Sub-group analyses

There are planned subgroups and interaction analyses proposed for this study. The a priori identified subgroups that will be used for the subgroup analyses are as follows:

i. Age
    - Quartiles

ii. Gender
    - Male versus Female
iii. Socio-economic status - Quintile of IMD 2015
    - 1 - Least deprived vs 2 vs 3 vs 4 vs 5 - Most deprived
iv. Duration of delirium
    - No delirium vs Delirium < median duration vs Delirium ≥ median duration
v. State trait anxiety inventory score (STAI)
    - Quartiles
vi. Surgical status
    - Emergency/urgent surgery vs Elective/scheduled surgery vs Non-surgical
vii. Overall site engagement (from process evaluation work)
    - Low vs Medium vs High
viii. Heterogeneity of treatment effect
    - Derivation of a risk prediction model for the primary outcome using the usual care patients' data adjusting for a priori important covariates (age, gender, socioeconomic status, duration of delirium, STAI, surgical status) and then grouping patients based on quintiles of predicted risk of outcome

The evaluation of the treatment effect on the primary outcome of this study will be carried out using a formal test of interaction which will be obtained from the linear mixed effect regression models.[24] The linear mixed effect model will contain a main effect term denoting the specific subgroup of interest, a main effect term for treatment group and a subgroup x treatment interaction term.

4.2.8. Process evaluation

Analysis of the process evaluation will use a combination of qualitative and quantitative methods to assess and describe the variation in the delivery of the intervention across sites. Analysis of the process evaluation will be conducted independent of the Trial team before the outcome evaluation to avoid any bias in the interpretation of the process data and to generate hypotheses that may be subsequently tested in statistical analyses of integrated process and outcome data. The structural mean models described above will be extended to incorporate additional potential mediator variables on the causal pathway between treatment allocation and treatment effect identified by the independent process evaluation team, e.g. nurse competence following training, adherence to the therapeutic approach, adherence to therapy and overall site engagement.[25]

4.2.9. Economic evaluation

A full CEA will be undertaken to assess the relative cost-effectiveness of psychological assessment followed by stress support sessions for those assessed as being at high risk of psychological morbidity versus usual care. Resource use and outcome data collected as part of the cluster-RCT will be used to report cost-effectiveness at six months and to project the lifetime cost-effectiveness of each strategy.

The cost analysis will take a health and personal health services perspective.[26] Cost will be calculated from patient level resource use data on length of stay in critical care and hospital, for the index admission and any readmission before six months (recorded in the trial dataset), use of personal health services after hospital discharge and within six months post-randomisation (collected through patient questionnaire), and additional staff time required to deliver the interventions (collected from site visits). Resource use data from the site visits, cluster-RCT dataset and six-month questionnaires will be combined with unit costs from the NHS Payment by Results database and from local Trust Finance Departments, to report the total costs per patient at six months for intervention versus usual care.[27,28]

HRQoL data from the EQ-5D-5L questionnaires at six months will be combined with survival data to report QALYs at six months. QALY will be calculated by valuing each patient's survival time by their health-related QOL at six months according to the 'area under the curve' approach.[29] For six month survivors, QALYs will be calculated using the EQ-5D scores at six months, assuming an EQ-5D score of zero at randomisation, and a linear interpolation between randomisation and six months. For decedents between randomisation and six months, we will assume zero QALYs.

The CEA will follow the intention-to-treat principle and report the mean (95% confidence interval) incremental costs, QALYs and net monetary benefit (NMB) at six months. Missing data in resource use and HRQoL will be handled with multiple imputation methods as described in the clinical analysis section. As a sensitivity analysis, Bayesian pattern-mixture models will be used to allow departures from MAR for the missing HRQoL, using a similar approach to that for the clinical effectiveness primary outcome.

The CEA will use GLMMs that allow for clustering of patients[30] including site as a random effect variable and period as a fixed effect variable. The analysis will adjust for pre-specified baseline covariates at both patient and site level. The primary effect estimate will be the interaction (difference in difference) between treatment groups and time period. The cost-effectiveness analysis will use this model to estimate the effect of the intervention on mean

cost and mean QALY (allowing for the correlation between the costs and QALY at the individual and cluster level).

Lifetime cost-effectiveness will be projected by summarising the relative effects of alternative strategies on long-term survival and HRQoL, informed by extrapolations of patient survival data.[31,32] The long-term modelling will extrapolate from the cluster-RCT data by fitting alternative parametric survival curves (e.g. Weibull, exponential, lognormal, log logistic and Gompertz) to the maximum available survival data recorded in the trial dataset. The chosen method of survival extrapolation for the base case analysis will be the one judged most plausible[33] according to model fit (Akaike information criteria (AIC) or Bayesian information criteria (BIC)), and in comparison with age-gender matched all-cause mortality. Quality of life generally deteriorates after critical care discharge for up to 6 months and then slowly improves over time but remains lower than that in the general population over long-term.[34] In the base case analysis, quality of life decrement of the study population compared with age-gender matched population[35] at six months will be applied allowing for improvement in quality of life over the years of excess mortality. After period of excess mortality, quality of life from age-gender matched general population will be applied. Lifetime costs attributable to initial episode of critical illness will be estimated by utilising longer term readmission costs data to patients who were randomised early. The longer term costs will be applied over the period of excess mortality. Predicted survival and HRQoL will be combined to report lifetime QALYs, and to project lifetime incremental costs, incremental QALYs, and incremental net benefits for the alternative strategies of care. Sensitivity analyses will test whether the results are robust to methodological assumptions (e.g. specification of the statistical model, extrapolation approach, alternative HRQoL assumptions, and learning curve effects).

Adherence adjusted analysis and subgroup analysis will be undertaken for the pre-specified subgroups as per the analysis of clinical effectiveness.

5. Reporting conventions

The following reporting conventions will be adopted for the SAP. These conventions will enhance the review of the study report and help to standardize presentation with common notations.

i. Sample sizes will be presented for each treatment group as totals in the column header as “(N = xxx)”, where appropriate.

ii. Sample sizes shown with summary statistics are the samples sizes (n) of patients with non-missing values.

iii. All summaries for categorical variables will include all categories that were available and will not be restricted to those with at least one response.

iv. Summaries for continuous variables that are approximately normally distributed will be reported as n, mean and standard deviation.

v. Summaries for continuous variables that are not normally distributed will be reported as n, median and quartiles.

vi. All percentages will be rounded and reported to a single decimal place (xx.x%). A percentage of 0% will be reported as “0%”; a percentage of 100% will be reported as “100%”.

vii. Summaries that include P-values will report the P-value to three decimal places with a leading zero (0.xxx). P-values of less than 0.0005 will be reported as “<0.001” not “0.000”.

viii. Missing values for both numeric and string variables will be presented as dashes (“---”) or as “Not available” / “Not applicable” / “Not reported” (as appropriate) in tables or data listings.

6. Proposed tables and figures

6.1. Clinical evaluation tables

Table 1: Baseline demographic and clinical variables by treatment groups

Variables	Baseline period		Intervention period	
	Intervention	Usual Care	Intervention	Usual Care
	N = XXX	N = XXX	N = XXX	N = XXX
Demography				
Age (years):				
Mean (SD)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)
Median (IQR)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)
Gender:				
Female, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Male, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Ethnicity:				
White, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Mixed, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Asian/Asian British, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Black/Black British, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Other, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Not stated, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Quintile of IMD 2015:				
1 - Least deprived, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
2, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
3, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
4, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
5 - Most deprived, n (%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Documented pre-existing anxiety/depression:				
Anxiety, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Depression, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Both, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
None, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Planned admission to the critical care unit following elective/schedule surgery				
Yes, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
No, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
ICNARC Physiology Score:				
Mean (SD)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)
Median (IQR)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)
APACHE II score:				
Mean (SD)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)
Median (IQR)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)

n: Number of patients; %: Percentage of patients; N: Total number of patients
SD: Standard deviation; IQR: Inter-quartile range; BMI: Body mass index.

Table 1: Con't

Variables	Baseline period		Intervention period	
	Intervention	Usual Care	Intervention	Usual Care
	N = XXX	N = XXX	N = XXX	N = XXX
Duration of critical care unit stay prior to consent:				
Mean (SD)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)
Median (IQR)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)
Number of days experiencing delirium in the critical care unit prior to consent				
Mean (SD)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)
Median (IQR)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)
Last NEWS prior to consent				
Mean (SD)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)
Median (IQR)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)
STAI-6 at time of consent				
Mean (SD)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)
Median (IQR)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)
HRQOL (health thermometer score) at time of consent:				
Mean (SD)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)	XX.X (XX.X)
Median (IQR)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)	XX (XX,XX)

Table 2: Concomitant medications used by treatment groups

Variables	Baseline period		Intervention period	
	Intervention	Usual Care	Intervention	Usual Care
	N = XXX	N = XXX	N = XXX	N = XXX
Sedatives/anxiolytics/anaesthetics received:				
Chlordiazepoxide, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Clobazam, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Clonidine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Desflurane, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Dexmedetomidine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Diazepam, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Etomidate, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Halothane, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Isoflurane, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Ketamine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Lorazepam, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Midazolam, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Propofol, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Sevoflurane, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Thiopentone, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Other, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Sleep medication received:				
Flurazepam, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Lormetazepam, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Nitrazepam, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Temazepam, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Zolpidem, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Zopiclone, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Other, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Benzodiazepines	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Antipsychotic medication received:				
Chlorpromazine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Clozapine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Flupentixol, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Haloperidol, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Olanzapine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Quetiapine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Risperidone, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Other, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)

Table 2: Con't

Variables	Baseline period		Intervention period	
	Intervention	Usual Care	Intervention	Usual Care
	N = XXX	N = XXX	N = XXX	N = XXX
Analgesics received:				
Alfentanil, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Co-codamol, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Codeine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Co-dydramol, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Diamorphine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Dihydrocodeine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Fentanyl, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Morphine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Oxycodone, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Remifentanil, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Tramadol, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Other, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Antidepressants received:				
Amitriptyline, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Citalopram, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Fluoxetine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Mirtazapine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Paroxetine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Reboxetine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Sertraline, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Venlafaxine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Other, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Actual vasoactive agent received:				
Adrenaline, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Dobutamine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Dopamine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Dopexamine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Metaraminol, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Noradrenaline, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Phenylephrine, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Vasopressin, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Other, n(%)	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)

n: number of patients; %: percentage of patients; N: total number of patients

Table 3: Linear mixed effect model for PSS-SR at six months – primary analysis

Variables	Coefficient	95% CI	P-value
Fixed effects at the site level*:			
Teaching status of hospital*:			
Teaching	0		
Non-Teaching	XX.X	XX.X , XX.X	0.XXX
Number of beds in the critical care unit (per additional bed)*	XX.X	XX.X , XX.X	0.XXX
Number of CCU admissions receiving Level 3 care staying at least 48hr during the pre-trial period,1 April 2014 to 31 March 2015 (per additional 100 admissions)*	XX.X	XX.X , XX.X	0.XXX
Allocated treatment group:			
Usual care	0		
Intervention	XX.X	XX.X , XX.X	0.XXX
Fixed effects at the patient level:			
Time period:			
Baseline period	0		
Intervention period	XX.X	XX.X , XX.X	0.XXX
Interaction between time period and treatment group:			
Intervention period * Intervention group	XX.X	XX.X , XX.X	0.XXX
Age in years (restricted cubic splines, 4 knots)			
Age spline 1	XX.X	XX.X , XX.X	0.XXX
Age spline 2	XX.X	XX.X , XX.X	
Age spline 3	XX.X	XX.X , XX.X	
Gender:			
Male	0		
Females	XX.X	XX.X , XX.X	0.XXX
Ethnicity:			0.XXX
White, n (%)	0		
Mixed, n (%)	XX.X	XX.X , XX.X	
Asian/Asian British, n (%)	XX.X	XX.X , XX.X	
Black/Black British, n (%)	XX.X	XX.X , XX.X	
Others, n (%)	XX.X	XX.X , XX.X	
Quintile of IMD 2015:			0.XXX
1 - Least deprived, n (%)	0		
2, n (%)	XX.X	XX.X , XX.X	
3, n (%)	XX.X	XX.X , XX.X	
4, n (%)	XX.X	XX.X , XX.X	
5 - Most deprived, n (%)	XX.X	XX.X , XX.X	
Pre-existing anxiety/depression:			0.XXX
Anxiety, n(%)	0		
Depression, n(%)	XX.X	XX.X , XX.X	
Both, n(%)	XX.X	XX.X , XX.X	
None, n(%)	XX.X	XX.X , XX.X	
Planned admission to the CCU following elective/schedule surgery			
Yes, n(%)	0		

No, n(%)	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score from the first 24h following admission to the critical care unit (restricted cubic splines, 4 knots)			
ICNARC Physiology Score spline 1	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score spline 2	XX.X	XX.X , XX.X	
ICNARC Physiology Score spline 3	XX.X	XX.X , XX.X	
Random Effects			
Site	XX.X	XX.X , XX.X	-

CI: Confidence interval.

* - Covariates used to balance treatment allocation

Table 4a: Linear mixed effect model for days alive and free from sedation to day 30

Variables	Coefficient	95% CI	P-value
Fixed effects at the site level*:			
Teaching status of hospital*:			
Teaching	0		
Non-Teaching	XX.X	XX.X , XX.X	0.XXX
Number of beds in the critical care unit (per additional bed)*	XX.X	XX.X , XX.X	0.XXX
Number of CCU admissions receiving Level 3 care staying at least 48hr during the pre-trial period,1 April 2014 to 31 March 2015 (per additional 100 admissions)*	XX.X	XX.X , XX.X	0.XXX
Allocated treatment group:			
Usual care	0		
Intervention	XX.X	XX.X , XX.X	0.XXX
Fixed effects at the patient level:			
Time period:			
Baseline period	0		
Intervention period	XX.X	XX.X , XX.X	0.XXX
Interaction between time period and treatment group:			
Intervention period * Intervention group	XX.X	XX.X , XX.X	0.XXX
Age in years (restricted cubic splines, 4 knots)			
Age spline 1	XX.X	XX.X , XX.X	0.XXX
Age spline 2	XX.X	XX.X , XX.X	
Age spline 3	XX.X	XX.X , XX.X	
Gender:			
Male	0		
Females	XX.X	XX.X , XX.X	0.XXX
Ethnicity:			0.XXX
White, n (%)	0		
Mixed, n (%)	XX.X	XX.X , XX.X	
Asian/Asian British, n (%)	XX.X	XX.X , XX.X	
Black/Black British, n (%)	XX.X	XX.X , XX.X	
Others, n (%)	XX.X	XX.X , XX.X	
Quintile of IMD 2015:			0.XXX
1 - Least deprived, n (%)	0		
2, n (%)	XX.X	XX.X , XX.X	
3, n (%)	XX.X	XX.X , XX.X	
4, n (%)	XX.X	XX.X , XX.X	
5 - Most deprived, n (%)	XX.X	XX.X , XX.X	
Pre-existing anxiety/depression:			0.XXX
Anxiety, n(%)	0		
Depression, n(%)	XX.X	XX.X , XX.X	
Both, n(%)	XX.X	XX.X , XX.X	
None, n(%)	XX.X	XX.X , XX.X	
Planned admission to the CCU following elective/schedule surgery			
Yes, n(%)	0		

No, n(%)	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score from the first 24h following admission to the critical care unit (restricted cubic splines, 4 knots)			
ICNARC Physiology Score spline 1	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score spline 2	XX.X	XX.X , XX.X	
ICNARC Physiology Score spline 3	XX.X	XX.X , XX.X	
Random Effects			
Site	XX.X	XX.X , XX.X	-

CI: Confidence interval.

* - Covariates used to balance treatment allocation

Table 4b: Linear mixed effect model for duration of critical care unit stay

Variables	Coefficient	95% CI	P-value
Fixed effects at the site level*:			
Teaching status of hospital*:			
Teaching	0		
Non-Teaching	XX.X	XX.X , XX.X	0.XXX
Number of beds in the critical care unit (per additional bed)*	XX.X	XX.X , XX.X	0.XXX
Number of CCU admissions receiving Level 3 care staying at least 48hr during the pre-trial period,1 April 2014 to 31 March 2015 (per additional 100 admissions)*	XX.X	XX.X , XX.X	0.XXX
Allocated treatment group:			
Usual care	0		
Intervention	XX.X	XX.X , XX.X	0.XXX
Fixed effects at the patient level:			
Time period:			
Baseline period	0		
Intervention period	XX.X	XX.X , XX.X	0.XXX
Interaction between time period and treatment group:			
Intervention period * Intervention group	XX.X	XX.X , XX.X	0.XXX
Age in years (restricted cubic splines, 4 knots)			
Age spline 1	XX.X	XX.X , XX.X	0.XXX
Age spline 2	XX.X	XX.X , XX.X	
Age spline 3	XX.X	XX.X , XX.X	
Gender:			
Male	0		
Females	XX.X	XX.X , XX.X	0.XXX
Ethnicity:			0.XXX
White, n (%)	0		
Mixed, n (%)	XX.X	XX.X , XX.X	
Asian/Asian British, n (%)	XX.X	XX.X , XX.X	
Black/Black British, n (%)	XX.X	XX.X , XX.X	
Others, n (%)	XX.X	XX.X , XX.X	
Quintile of IMD 2015:			0.XXX
1 - Least deprived, n (%)	0		
2, n (%)	XX.X	XX.X , XX.X	
3, n (%)	XX.X	XX.X , XX.X	
4, n (%)	XX.X	XX.X , XX.X	
5 - Most deprived, n (%)	XX.X	XX.X , XX.X	
Pre-existing anxiety/depression:			0.XXX
Anxiety, n(%)	0		
Depression, n(%)	XX.X	XX.X , XX.X	
Both, n(%)	XX.X	XX.X , XX.X	
None, n(%)	XX.X	XX.X , XX.X	
Planned admission to the CCU following elective/schedule surgery			
Yes, n(%)	0		

No, n(%)	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score from the first 24h following admission to the critical care unit (restricted cubic splines, 4 knots)			
ICNARC Physiology Score spline 1	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score spline 2	XX.X	XX.X , XX.X	
ICNARC Physiology Score spline 3	XX.X	XX.X , XX.X	
Random Effects			
Site	XX.X	XX.X , XX.X	-

CI: Confidence interval.

* - Covariates used to balance treatment allocation

Table 4c: Logistic mixed effect model for PSS-SR greater than 18 points at six months

Variables	Odds ratio	95% CI	P-value
Fixed effects at the site level*:			
Teaching status of hospital*:			
Teaching	0		
Non-Teaching	XX.X	XX.X , XX.X	0.XXX
Number of beds in the critical care unit (per additional bed)*	XX.X	XX.X , XX.X	0.XXX
Number of CCU admissions receiving Level 3 care staying at least 48hr during the pre-trial period,1 April 2014 to 31 March 2015 (per additional 100 admissions)*	XX.X	XX.X , XX.X	0.XXX
Allocated treatment group:			
Usual care	0		
Intervention	XX.X	XX.X , XX.X	0.XXX
Fixed effects at the patient level:			
Time period:			
Baseline period	0		
Intervention period	XX.X	XX.X , XX.X	0.XXX
Interaction between time period and treatment group:			
Intervention period * Intervention group	XX.X	XX.X , XX.X	0.XXX
Age in years (restricted cubic splines, 4 knots)			
Age spline 1	XX.X	XX.X , XX.X	0.XXX
Age spline 2	XX.X	XX.X , XX.X	
Age spline 3	XX.X	XX.X , XX.X	
Gender:			
Male	0		
Females	XX.X	XX.X , XX.X	0.XXX
Ethnicity:			0.XXX
White, n (%)	0		
Mixed, n (%)	XX.X	XX.X , XX.X	
Asian/Asian British, n (%)	XX.X	XX.X , XX.X	
Black/Black British, n (%)	XX.X	XX.X , XX.X	
Others, n (%)	XX.X	XX.X , XX.X	
Quintile of IMD 2015:			0.XXX
1 - Least deprived, n (%)	0		
2, n (%)	XX.X	XX.X , XX.X	
3, n (%)	XX.X	XX.X , XX.X	
4, n (%)	XX.X	XX.X , XX.X	
5 - Most deprived, n (%)	XX.X	XX.X , XX.X	
Pre-existing anxiety/depression:			0.XXX
Anxiety, n(%)	0		
Depression, n(%)	XX.X	XX.X , XX.X	
Both, n(%)	XX.X	XX.X , XX.X	
None, n(%)	XX.X	XX.X , XX.X	
Planned admission to the CCU following elective/schedule surgery			
Yes, n(%)	0		

No, n(%)	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score from the first 24h following admission to the critical care unit (restricted cubic splines, 4 knots)			
ICNARC Physiology Score spline 1	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score spline 2	XX.X	XX.X , XX.X	
ICNARC Physiology Score spline 3	XX.X	XX.X , XX.X	
Random Effects			
Site	XX.X	XX.X , XX.X	-

CI: Confidence interval.

* - Covariates used to balance treatment allocation

Table 4d: Linear mixed effect model for HADS depression score at six month

Variables	Coefficient	95% CI	P-value
Fixed effects at the site level*:			
Teaching status of hospital*:			
Teaching	0		
Non-Teaching	XX.X	XX.X , XX.X	0.XXX
Number of beds in the critical care unit (per additional bed)*	XX.X	XX.X , XX.X	0.XXX
Number of CCU admissions receiving Level 3 care staying at least 48hr during the pre-trial period,1 April 2014 to 31 March 2015 (per additional 100 admissions)*	XX.X	XX.X , XX.X	0.XXX
Allocated treatment group:			
Usual care	0		
Intervention	XX.X	XX.X , XX.X	0.XXX
Fixed effects at the patient level:			
Time period:			
Baseline period	0		
Intervention period	XX.X	XX.X , XX.X	0.XXX
Interaction between time period and treatment group:			
Intervention period * Intervention group	XX.X	XX.X , XX.X	0.XXX
Age in years (restricted cubic splines, 4 knots)			
Age spline 1	XX.X	XX.X , XX.X	0.XXX
Age spline 2	XX.X	XX.X , XX.X	
Age spline 3	XX.X	XX.X , XX.X	
Gender:			
Male	0		
Females	XX.X	XX.X , XX.X	0.XXX
Ethnicity:			0.XXX
White, n (%)	0		
Mixed, n (%)	XX.X	XX.X , XX.X	
Asian/Asian British, n (%)	XX.X	XX.X , XX.X	
Black/Black British, n (%)	XX.X	XX.X , XX.X	
Others, n (%)	XX.X	XX.X , XX.X	
Quintile of IMD 2015:			0.XXX
1 - Least deprived, n (%)	0		
2, n (%)	XX.X	XX.X , XX.X	
3, n (%)	XX.X	XX.X , XX.X	
4, n (%)	XX.X	XX.X , XX.X	
5 - Most deprived, n (%)	XX.X	XX.X , XX.X	
Pre-existing anxiety/depression:			0.XXX
Anxiety, n(%)	0		
Depression, n(%)	XX.X	XX.X , XX.X	
Both, n(%)	XX.X	XX.X , XX.X	
None, n(%)	XX.X	XX.X , XX.X	
Planned admission to the CCU following elective/schedule surgery			
Yes, n(%)	0		

No, n(%)	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score from the first 24h following admission to the critical care unit (restricted cubic splines, 4 knots)			
ICNARC Physiology Score spline 1	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score spline 2	XX.X	XX.X , XX.X	
ICNARC Physiology Score spline 3	XX.X	XX.X , XX.X	
Random Effects			
Site	XX.X	XX.X , XX.X	-

CI: Confidence interval.

* - Covariates used to balance treatment allocation

Table 4e: Linear mixed effect model for HADS anxiety score at six months

Variables	Coefficient	95% CI	P-value
Fixed effects at the site level*:			
Teaching status of hospital*:			
Teaching	0		
Non-Teaching	XX.X	XX.X , XX.X	0.XXX
Number of beds in the critical care unit (per additional bed)*	XX.X	XX.X , XX.X	0.XXX
Number of CCU admissions receiving Level 3 care staying at least 48hr during the pre-trial period,1 April 2014 to 31 March 2015 (per additional 100 admissions)*	XX.X	XX.X , XX.X	0.XXX
Allocated treatment group:			
Usual care	0		
Intervention	XX.X	XX.X , XX.X	0.XXX
Fixed effects at the patient level:			
Time period:			
Baseline period	0		
Intervention period	XX.X	XX.X , XX.X	0.XXX
Interaction between time period and treatment group:			
Intervention period * Intervention group	XX.X	XX.X , XX.X	0.XXX
Age in years (restricted cubic splines, 4 knots)			
Age spline 1	XX.X	XX.X , XX.X	0.XXX
Age spline 2	XX.X	XX.X , XX.X	
Age spline 3	XX.X	XX.X , XX.X	
Gender:			
Male	0		
Females	XX.X	XX.X , XX.X	0.XXX
Ethnicity:			0.XXX
White, n (%)	0		
Mixed, n (%)	XX.X	XX.X , XX.X	
Asian/Asian British, n (%)	XX.X	XX.X , XX.X	
Black/Black British, n (%)	XX.X	XX.X , XX.X	
Others, n (%)	XX.X	XX.X , XX.X	
Quintile of IMD 2015:			0.XXX
1 - Least deprived, n (%)	0		
2, n (%)	XX.X	XX.X , XX.X	
3, n (%)	XX.X	XX.X , XX.X	
4, n (%)	XX.X	XX.X , XX.X	
5 - Most deprived, n (%)	XX.X	XX.X , XX.X	
Pre-existing anxiety/depression:			0.XXX
Anxiety, n(%)	0		
Depression, n(%)	XX.X	XX.X , XX.X	
Both, n(%)	XX.X	XX.X , XX.X	
None, n(%)	XX.X	XX.X , XX.X	
Planned admission to the CCU following elective/schedule surgery			
Yes, n(%)	0		

No, n(%)	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score from the first 24h following admission to the critical care unit (restricted cubic splines, 4 knots)			
ICNARC Physiology Score spline 1	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score spline 2	XX.X	XX.X , XX.X	
ICNARC Physiology Score spline 3	XX.X	XX.X , XX.X	
Random Effects			
Site	XX.X	XX.X , XX.X	-

CI: Confidence interval.

* - Covariates used to balance treatment allocation

Table 4f: Linear mixed effect model for health related quality of life at six months

Variables	Coefficient	95% CI	P-value
Fixed effects at the site level*:			
Teaching status of hospital*:			
Teaching	0		
Non-Teaching	XX.X	XX.X , XX.X	0.XXX
Number of beds in the critical care unit (per additional bed)*	XX.X	XX.X , XX.X	0.XXX
Number of CCU admissions receiving Level 3 care staying at least 48hr during the pre-trial period,1 April 2014 to 31 March 2015 (per additional 100 admissions)*	XX.X	XX.X , XX.X	0.XXX
Allocated treatment group:			
Usual care	0		
Intervention	XX.X	XX.X , XX.X	0.XXX
Fixed effects at the patient level:			
Time period:			
Baseline period	0		
Intervention period	XX.X	XX.X , XX.X	0.XXX
Interaction between time period and treatment group:			
Intervention period * Intervention group	XX.X	XX.X , XX.X	0.XXX
Age in years (restricted cubic splines, 4 knots)			
Age spline 1	XX.X	XX.X , XX.X	0.XXX
Age spline 2	XX.X	XX.X , XX.X	
Age spline 3	XX.X	XX.X , XX.X	
Gender:			
Male	0		
Females	XX.X	XX.X , XX.X	0.XXX
Ethnicity:			0.XXX
White, n (%)	0		
Mixed, n (%)	XX.X	XX.X , XX.X	
Asian/Asian British, n (%)	XX.X	XX.X , XX.X	
Black/Black British, n (%)	XX.X	XX.X , XX.X	
Others, n (%)	XX.X	XX.X , XX.X	
Quintile of IMD 2015:			0.XXX
1 - Least deprived, n (%)	0		
2, n (%)	XX.X	XX.X , XX.X	
3, n (%)	XX.X	XX.X , XX.X	
4, n (%)	XX.X	XX.X , XX.X	
5 - Most deprived, n (%)	XX.X	XX.X , XX.X	
Pre-existing anxiety/depression:			0.XXX
Anxiety, n(%)	0		
Depression, n(%)	XX.X	XX.X , XX.X	
Both, n(%)	XX.X	XX.X , XX.X	
None, n(%)	XX.X	XX.X , XX.X	
Planned admission to the CCU following elective/schedule surgery			
Yes, n(%)	0		

No, n(%)	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score from the first 24h following admission to the critical care unit (restricted cubic splines, 4 knots)			
ICNARC Physiology Score spline 1	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score spline 2	XX.X	XX.X , XX.X	
ICNARC Physiology Score spline 3	XX.X	XX.X , XX.X	
Random Effects			
Site	XX.X	XX.X , XX.X	-

CI: Confidence interval.

* - Covariates used to balance treatment allocation

Table 5: Structural mean models for PSS-SR at six months using randomised allocated treatment as an instrumental variable

Variables	Coefficient	95% CI	P-value
Fixed effects at the site level*:			
Teaching status of hospital*:			
Teaching	0		
Non-Teaching	XX.X	XX.X , XX.X	0.XXX
Number of beds in the critical care unit (per additional bed)*	XX.X	XX.X , XX.X	0.XXX
Number of CCU admissions receiving Level 3 care staying at least 48hr during the pre-trial period,1 April 2014 to 31 March 2015 (per additional 100 admissions)*	XX.X	XX.X , XX.X	0.XXX
Allocated treatment group:			
Usual care	0		
Intervention	XX.X	XX.X , XX.X	0.XXX
Fixed effects at the patient level:			
Time period:			
Baseline period	0		
Intervention period	XX.X	XX.X , XX.X	0.XXX
Interaction between time period and treatment group:			
Intervention period * Intervention group	XX.X	XX.X , XX.X	0.XXX
Age in years (restricted cubic splines, 4 knots)			
Age spline 1	XX.X	XX.X , XX.X	0.XXX
Age spline 2	XX.X	XX.X , XX.X	
Age spline 3	XX.X	XX.X , XX.X	
Gender:			
Male	0		
Females	XX.X	XX.X , XX.X	0.XXX
Ethnicity:			0.XXX
White, n (%)	0		
Mixed, n (%)	XX.X	XX.X , XX.X	
Asian/Asian British, n (%)	XX.X	XX.X , XX.X	
Black/Black British, n (%)	XX.X	XX.X , XX.X	
Others, n (%)	XX.X	XX.X , XX.X	
Quintile of IMD 2015:			0.XXX
1 - Least deprived, n (%)	0		
2, n (%)	XX.X	XX.X , XX.X	
3, n (%)	XX.X	XX.X , XX.X	
4, n (%)	XX.X	XX.X , XX.X	
5 - Most deprived, n (%)	XX.X	XX.X , XX.X	
Pre-existing anxiety/depression:			0.XXX
Anxiety, n(%)	0		
Depression, n(%)	XX.X	XX.X , XX.X	
Both, n(%)	XX.X	XX.X , XX.X	
None, n(%)	XX.X	XX.X , XX.X	
Planned admission to the CCU following elective/schedule surgery			
Yes, n(%)	0		

No, n(%)	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score from the first 24h following admission to the critical care unit (restricted cubic splines, 4 knots)			
ICNARC Physiology Score spline 1	XX.X	XX.X , XX.X	0.XXX
ICNARC Physiology Score spline 2	XX.X	XX.X , XX.X	
ICNARC Physiology Score spline 3	XX.X	XX.X , XX.X	
Random Effects			
Site	XX.X	XX.X , XX.X	-

6.2. Economic evaluation tables

Table 6: Parameter estimates of the parametric survival models used for extrapolating survival curves

Distribution	Parameter estimates	
	Scale/Rate	Shape
Exponential	xx.x	N/A
Weibull	xx.x	xx.x
Lognormal(sdlog/meanlog)	xx.x	xx.x
Log-logistic	xx.x	xx.x
Gompertz	xx.x	xx.x

Table 7: Survival probabilities of the parametric survival models

Time (Years)	Exponential	Weibull	Lognormal	Log-logistic	Gompertz
0	1	1	1	1	1
1	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
2	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
3	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
.	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
.	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
.	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
.	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
.	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
.	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
.	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
.	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
.	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
.	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
98	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
99	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
100	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx

Table 8: Rank of Goodness of fit estimates (AIC and BIC) for parametric survival models

Distribution	AIC	BIC	Ranking
Exponential	xxx.x	xxx.x	x
Weibull	xxx.x	xxx.x	x
Lognormal	xxx.x	xxx.x	x
Log-logistic	xxx.x	xxx.x	x
Gompertz	xxx.x	xxx.x	x

6.3. Figures

Figure 1: CONSORT flow diagram

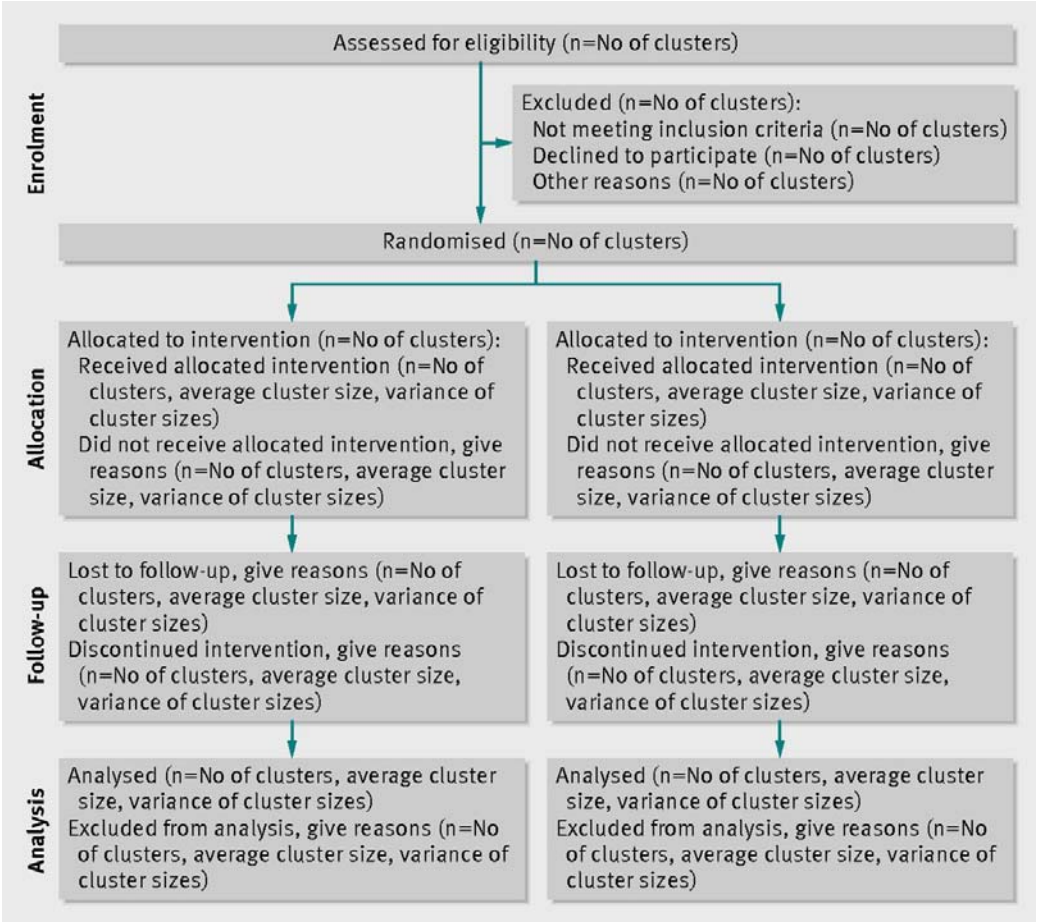


Figure 2: Kaplan-Meier plot of comparing intervention and control groups during intervention period

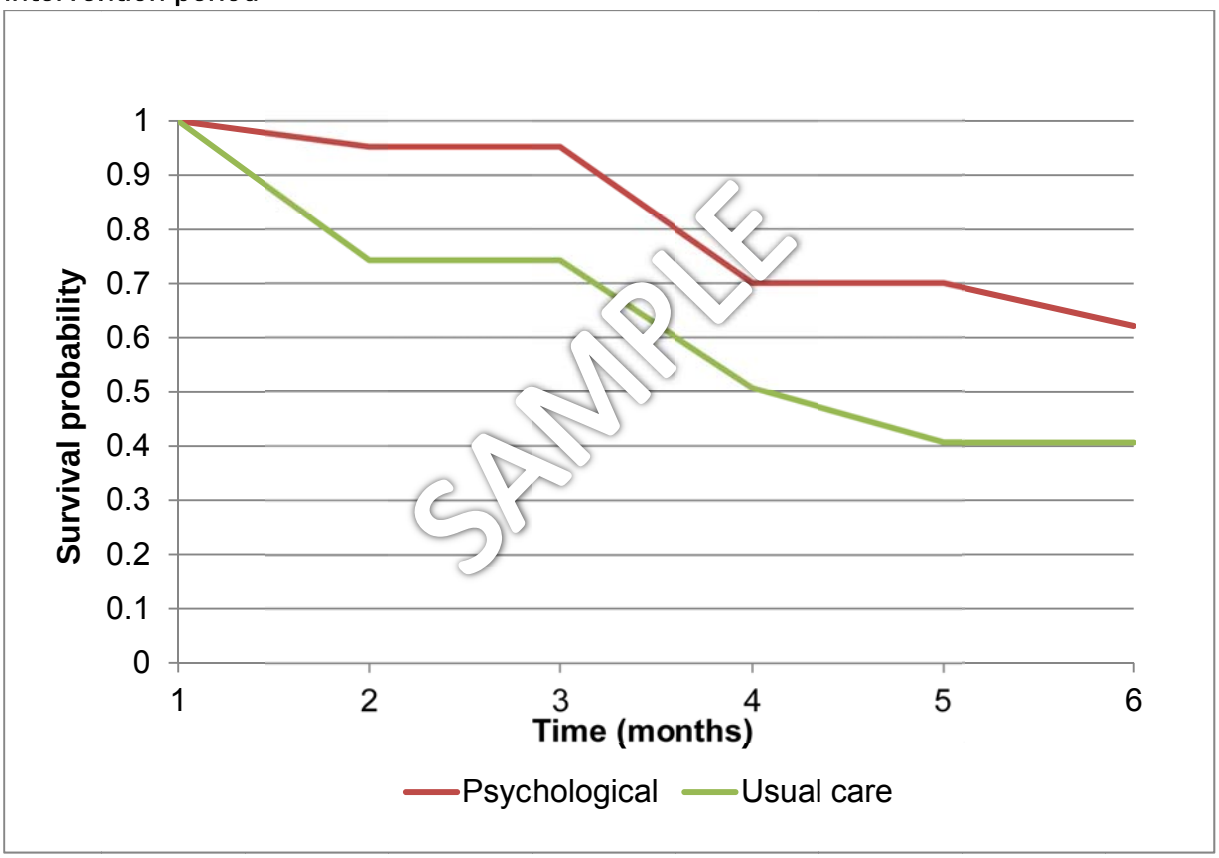
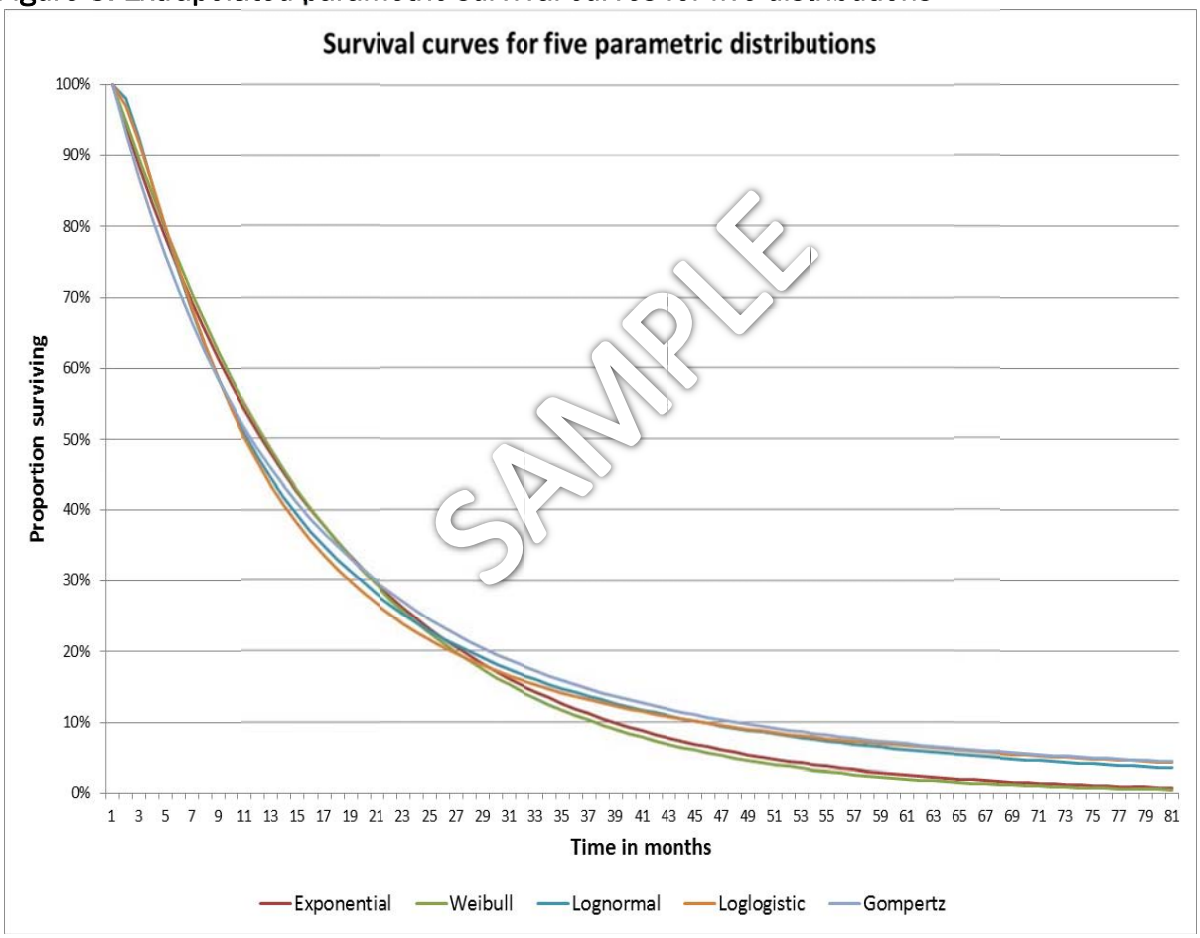


Figure 3: Extrapolated parametric survival curves for five distributions



7. References

1. Hatch R, McKechnie S, Griffiths J. Psychological intervention to prevent ICU-related PTSD: who, when and for how long? Crit Care 2011; 15:141.
2. Ely EW, Inouye SK, Bernard GR, et al. Delirium in mechanically ventilated patients: validity and reliability of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). J Am Med Assoc 2001; 286:2703-10.
3. Sessler CN, Gosnell MS, Grap MJ, et al. The Richmond Agitation-Sedation Scale: validity and reliability in adult intensive care unit patients. Am J Respir Crit Care Med 2002; 166:1338-44.
4. Foa EB, Cashman L, Jaycox L, et al. The validation of a self-report measure of posttraumatic stress disorder: the Posttraumatic Diagnostic Scale. Psychol Assess 1997; 9:445.
5. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983; 67:361-70.
6. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin trials 2007; 28:182-91.
7. Wade DM, Howell DC, Weinman JA, et al. Investigating risk factors for psychological morbidity three months after intensive care: a prospective cohort study. Crit Care 2012; 16:R192.
8. Hayes RJ, Bennett S. Simple sample size calculation for cluster-randomized trials. Int J Epidemiol 1999; 28:319-26.
9. Brewin CR, Fuchkan N, Huntley Z, et al. Outreach and screening following the 2005 London bombings: usage and outcomes. Psychol Med 2010; 40:2049-57.
10. Wade DM, Hankins M, Smyth DA, et al. Detecting acute distress and risk of future psychological morbidity in critically ill patients: validation of the intensive care psychological assessment tool. Crit Care 2014; 18:519.
11. Carter BR, Hood K. Balance algorithm for cluster randomized trials. BMC Med Res Methodol 2008; 8:65.
12. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/ (2013, accessed 4th May 2017).
13. Carpenter JR, Kenward MG. Multiple Imputation and its Application. Wiley, 2013.
14. Quartagno M, Carpenter JR. Multiple imputation for IPD meta-analysis: allowing for heterogeneity and studies with missing covariates. Stat Med 2016; 35:2938-54.
15. Department for Communities and Local Government. English indices of deprivation 2015. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015 (2015, accessed 4th May 2017).
16. Harrison DA, Parry GJ, Carpenter JR, et al. A new risk prediction model for critical care: the Intensive Care National Audit & Research Centre (ICNARC) model. Crit Care Med 2007; 35:1091-8.
17. Royal College of Physicians. National Early Warning Score (NEWS): Standardising the assessment of acute illness severity in the NHS. https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news (2012, accessed 4th May 2017).

18. Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). Br J Clin Psychol 1992; 3:301-6.

19. Molenberghs G, Fitzmaurice G, Kenward MG, et al. Handbook of Missing Data Methodology. Boca Raton, FL: Chapman & Hall/CRC, 2015.

20. Mason AJ, Gomes M, Grieve R, et al. Development of a practical approach to expert elicitation for randomised controlled trials with missing health outcomes: Application to the IMPROVE trial. Clin Trials 2017; 14:357-67.

21. Campbell MK, Piaggio G, Elbourne DR, et al. Consort 2010 statement: extension to cluster randomised trials. BMJ 2012; 345:e5661.

22. Draper, D. Rank-based robust analysis of linear models. I. Exposition and review. Stat Sci 1988; 3:239-57.

23. Maracy M, Dunn G. Estimating dose-response effects in psychological treatment trials: the role of instrumental variables. Stat Meth Med Res 2011; 20:191-215.

24. Wang R, Lagakos SW, Ware JH, et al. Statistics in medicine—reporting of subgroup analyses in clinical trials. N Engl J Med 2007; 357:2189-94.

25. Emsley R, Dunn G, White IR. Mediation and moderation of treatment effects in randomised controlled trials of complex interventions. Stat Meth Med Res 2010; 19:237-70.

26. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. London: National Institute for Health and Care Excellence, 2013.

27. Curtis L. Unit costs of health & social care 2011. Canterbury, UK: Personal Service Research Unit, University of Kent, 2011.

28. Department of Health. NHS reference costs 2013-14. London: Department of Health, 2014.

29. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Econ 2005; 14:487-96.

30. Grieve R, Nixon R, Thompson SG. Bayesian hierarchical models for cost-effectiveness analyses that use data from cluster randomized trials. Med Decis Making 2010; 30:163-75.

31. Wright JC, Plenderleith L, Ridley SA. Long-term survival following intensive care: subgroup analysis and comparison with the general population. Anaesth 2003; 58:637-42.

32. Green C, Dinnes J, Takeda AL, Cuthbertson BH. Evaluation of the cost-effectiveness of drotrecogin alfa (activated) for the treatment of severe sepsis in the United Kingdom. Int J Technol Assess Health Care 2006; 22:90-100.

33. Latimer NR. Survival analysis for economic evaluations alongside clinical trials—extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. Med Decis Making 2013; DOI: 10.1177/0272989x12472398.

34. Dowdy DW, Eid MP, Sedrakyan A, et al. Quality of life in adult survivors of critical illness: a systematic review of the literature. Intensive Care Med 2005; 31:611-20.

35. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value Health 2010; 13:509-18.