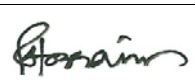
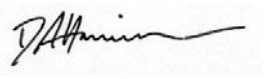
1 SUPPLEMENT 2: STATISTICIAL ANALYSIS PLAN

Trial:	**Study into the Reversal of Septic Shock with landiolol (Beta Blockade): STRESS-L Study**		
Chief Investigator:	**Dr Tony Whitehouse**		
Document: D Protocol / amendment* D PIS/ amendment* D Consent form / amendment* D CRF / amendment* D Publication(s) *(add details)* D Risk Assessment D Monitoring Plan D Statistical Analysis Plan		**Version: 1.0**	**Date:**

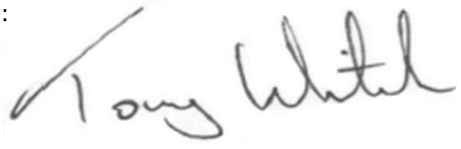
Reviewers:

By signing this form, you confirm the content of the document specified has been reviewed and found to be correct.

Name	**Role in trial**	**Signature**	**Date**
Dr Dipesh Mistry	*Statistician*		
Dr Anower Hossain	*Statistician*		
Professor Ranjit Lall	*Senior Statistician*		13/05/2022
Professor David Harrison	*DMC chair*		24/03/2021

Chief Investigator Approval:

I confirm that the document specified has been reviewed and I approve the use of this version.

Signature: 	Date:25 March 2022

Statistical Analysis Plan for the STRESS-L (Study into the Reversal of Septic Shock with landiolol (Beta Blockade)) Randomized Trial

Dipesh Mistry[1], Anower Hossain[2], Ranjit Lall[2], Emma Skilton[2], Belinder Ghuman[2], Gavin Perkins[2,3], Anthony Gordon[4], Julian Bion[5], Duncan Young[6], Danny McAuley[7,8], Mervyn Singer[9], Janet Lord[10], Simon Gates[11], Scott Regan[2], Tony Whitehouse[5,10]

1. Statistics and Decision Sciences, Janssen Pharmaceuticals R & D, High Wycombe, UK
2. Warwick Clinical Trials Unit, University of Warwick, Coventry, UK
3. Department of Critical Care, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK
4. Division of Anaesthetics, Pain Medicine & Intensive Care, Imperial College London, London, UK
5. Institute of Clinical Sciences, College of Medical and Dental Sciences, University of Birmingham, UK
6. Kadoorie Centre for Critical Care Research, Nuffield Division of Anaesthesia, University of Oxford, Oxford, UK
7. Regional Intensive Care Unit, Royal Victoria Hospital, Belfast Health and Social Care Trust, Belfast, UK
8. The Wellcome Wolfson Institute for Experimental Medicine, Queens University Belfast, Belfast, UK
9. Centre for Intensive Care Medicine, Department of Medicine and Wolfson Institute for Biomedical Research, University College, London, UK
10. Institute of Inflammation and Ageing, Birmingham University Research labs., Queen Elizabeth Hospital, Birmingham, UK
11. Cancer Clinical Trials Unit, University of Birmingham, Birmingham, UK

Corresponding author: Dr Ranjit Lall, Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick. Email: R.Lall@Warwick.ac.uk

Introduction

Background and rationale

In critically ill patients, sepsis is a major cause of mortality and morbidity, and its incidence is increasing.[1] The mortality rate from its most severe manifestation, septic shock, remains very high.[2] A single center randomized trial in Rome reported the use of esmolol in patients with septic shock and tachycardia requiring vasopressor therapy for more than 24 hours. Although this study was powered inadequately, marked improvements were observed in mortality and other secondary clinical outcomes thus warranting further investigation.[3]

The STRESS-L trial is a randomized controlled trial to assess the clinical effectiveness of landiolol infusion to reduce the heart rate between 80-94 bpm in comparison with standard care. landiolol was the selected beta-blocker rather than esmolol due to its superior beta-1 specificity and very short half-life making titration of heart rate easier. The half-life would also ease reversibility should severe side-effects occur. A more detailed explanation of the background, rationale, intervention, and trial design can be found in the protocol paper. A brief overview of the trial is presented in this paper.

This paper presents the statistical analysis plan (as purposed and agreed with the Data Monitoring Committee) for the STRESS-L trial which has been developed in line with published guidelines[4].

Objectives

The main objectives of this trial are to assess the efficacy, tolerability, safety, and mechanism of landiolol in patients with septic shock requiring prolonged (>24 hours) support with high-dose vasopressor agents.

Study methods

Trial design

The STRESS-L trial is a multi-center, open-label, non-blinded phase IIb randomized controlled trial (RCT). Participants are randomly allocated using a 1:1 ratio to either standard care (control group) or standard care and landiolol (intervention group). A computerized minimization randomization system was used, created by the Warwick Clinical Trials unit, to randomize participants. Randomization was minimized on recruiting site and noradrenaline dose (≥0.1 mcg/kg/min - 0.3 mcg/kg/min and >0.3 mcg/kg/min) where the dose reflects the participants severity.

Sample size

The primary outcome is the mean Sequential Organ Failure Assessment (SOFA) score over the first 14 days in ICU. Preliminary data on 324 patients from University Hospitals Birmingham satisfying the trial eligibility criteria suggested that the mean SOFA score over the first 14 days in ICU was 6.3, with standard deviation (SD) 2.4. Assuming a conservative SD of 2.8, data on 330 patients would be required to provide 90% power to detect a difference of 1 point between the landiolol and standard care groups, using a significance level of 0.05 (two-sided). To allow for 3% withdrawals and losses, the proposed sample size is 340. Although there were no formal interim analyses built into the study design, the sample size parameters, such as the overall standard deviation was monitored by the Data Monitoring Committee.

Timing of final analysis

The end of the trial is defined when the last recruited participant has completed their 90-day follow-up. Once all follow-up data has been entered onto the database, the data will be fully validated and cleaned after which the database will be locked, and the final analyses will be undertaken.

Timing of outcome assessments

All outcome assessments during the trial will be taken relative to the randomization date of the participant. Outcome assessments will be made on a daily basis from randomization to day 14 for all participants alive whilst in ICU. The mortality status of participants will also be collected at 28 days and 90 days post-randomization. Any serious adverse events (SAEs) will also be reported up to 90 days post-randomization.

Statistical principles

Confidence intervals and P values

All statistical tests will test for superiority and will be two-sided using a $P<0.05$ significance threshold. Estimates of effect size will be reported with 95% confidence intervals to express the statistical uncertainty. Adjustments for multiplicity will not be applied. Pre-specified subgroup analysis results will be interpreted with caution.

Adherence, protocol deviations and protocol violations

Adherence with the drug infusion protocol will be closely monitored. For all participants, the heart rate is monitored hourly from randomization to day 2, and then every 6 hours thereafter up to day 14. At each of these time intervals, the landiolol dose is adjusted according to the study drug infusion protocol depending on the heart rate. A participant is said to have not adhered (i.e., non-

complier) to the drug infusion protocol at a given interval if, any of these were satisfied, at any one point in their treatment phase:

1. The starting landiolol infusion dose is not 1.0 mcg/kg/min
2. The heart rate is <80 bpm and the landiolol infusion is not reduced by at least 1.0 mcg/kg/min
3. The heart rate is >94 bpm and the landiolol infusion is not increased by 1.0 mcg/kg/min
4. landiolol dose is not an integer
5. landiolol dose over 40mcg/kg/min (max dose)
6. landiolol restarted, 12 hours or more since last landiolol dose
7. On trial, landiolol not started

In addition, on a regular basis we assessed, a rate of non-adherence (i.e. number of non-compliant intervals out of the total number of intervals) of over 10% for a participant will trigger further investigation. Non-adherence at the site (i.e. mean rate of non-adherence for all participants within each site) of more than 15% will warrant further investigation at the site. At the end of the study, the rate of adherence in the intervention group will be summarized both at the participant level and site level using descriptive statistics (mean and standard deviation or median and interquartile range). The total number of protocol deviations and protocol violations will be summarized by treatment arm as well as the number and percentage of participants reporting them.

Bias assessment

If a participant is randomized to the landiolol plus standard treatment arm, ideally the landiolol should be dispensed, and the infusion started within one hour post randomization. Data collection will commence as soon as randomization occurs and will not be delayed until the landiolol is administered. As the landiolol arm requires additional data collection, the quality of data collected from randomization to the commencement of landiolol as well as after administering landiolol will be monitored for potential data collection bias during the study. More specifically the following will be monitored:

- <u>Data completeness up to the point of administering landiolol</u> - The proportion of data completeness will be summarized and compared between both arms for the SOFA score data, in/out fluids data and additional assessments data on day 0 as these outcomes are collected daily. Cardiovascular data is collected hourly and so the first 2 hours will be compared as by this time the landiolol would have been administered.

- Data completeness after administering landiolol - The proportion of data completeness will be summarized and compared between both arms for the SOFA score data, in/out fluids data and additional assessments data on day 1. The Cardiovascular data is collected hourly and so the hour 3 data will be compared for completeness between arms.

Analysis populations

The primary analysis method will be intention-to-treat, where participants will be analyzed according to the treatment, they were randomized to regardless of the treatment they actually received. All participants will be included in the analysis, irrespective of whether they adhered to the protocol.

Trial population

Screening data

A detailed summary of the screening data will be presented to describe the representativeness of the trial sample. This will include the following:

- Duration of screening (days)
- Total number of screened patients
- Number and percentage of eligible patients of those screened
- Number and percentage of randomized patients of those eligible summarizing reasons for non-recruitment
- Mean/median recruitment rate per month overall and by site

Eligibility

The STRESS-L trial protocol provides full details of the eligibility criteria (inclusion/exclusion criteria). Of those patients screened who are ineligible, the number and percentage will be summarized for each of the inclusion criteria that were not satisfied and/or the exclusion criteria that were met.

Recruitment

A CONSORT5 diagram will be used to illustrate the flow of participants throughout the trial. This will detail how many patients were:

- Assessed for eligibility
- Eligible for participation in the trial at screening
- Ineligible for participants in the trial at screening (stating reasons)
- Eligible for participation and were randomized
- Eligible for participants but were not randomized (stating reasons)
- Lost to follow-up at each follow-up time-point (stating reasons)
- Included in the final analyses at the primary endpoint listing reasons why participants were excluded

Withdrawal and follow-up

A participant's consent for the study may be withdrawn at any time by the participant themselves, a personal legal representative, a professional legal representative, or the treatment clinician. They can either withdraw from the intervention alone but remain on follow-up or they can withdraw completely i.e., no further participation in the trial. The number and percentage of participants who withdraw consent to further participation will be summarized by each group. Data will cease to be collected from the point of withdrawal, however data collected prior to withdrawal will be included in the final analyses. The decision to withdraw can be made by the participant, personal legal representative, professional legal representative, or clinician. This will also be summarized by group.

The number and percentage of participants assessed for follow-up will be calculated at the 28-day and 90-day follow-up time points as the number of participants assessed out of all randomized participants. The number and percentage of participants not assessed (i.e., due to death, withdrawal or lost to follow-up) at each of the follow-up time points will also be summarized in a similar manner out of the total number randomized. The follow-up completion rate will also be reported out of those expected to complete follow-up i.e. excluding those who died, and summarized by treatment arm.

Baseline patient characteristics

The baseline demographic characteristics and pre-randomization clinical measures of all randomized participants will be summarized by treatment group, but no formal statistical comparisons will be undertaken. In addition to these data, important process data (e.g., time from hospital admission to randomization) will also be summarized. Continuous variables will be summarized as mean (SD) or median and interquartile range (IQR). Counts and percentages will be used to summarize categorical variables. Below is a list of the demographic, clinical measures and process variable data that will be collected and summarized:

- Age (years)
- Gender (male/female)
- Weight (kg)
- Concomitant illness (yes/no)
- Main presumed/known site of infection at admission to ICU
  (Lungs/urine/abdomen/blood/other)
- Electrocardiogram (ECG) performed (yes/no)
- Chest x-ray taken (yes/no)
- Diffuse bilateral pulmonary infiltrates on x-ray (yes/no)
- Acute respiratory distress syndrome (yes/no)
- Pregnancy test done (yes/no – at investigators discretion)
- Steroid (mg) at randomization
- Beta blocker usage prior to randomization (yes/no)
- Baseline cardiovascular:
  - Arterial PaO2/PaCO2 (kPa)
  - Venous PaO2/PaCO2 (kPa)
  - Cardiac output (L/min)
  - Stroke volume (mL)
  - Mean arterial pressure (mmHg)
  - Heart rate (beats/min)
  - Atrial fibrillation (yes/no)
  - Noradrenaline dose (mcg/kg/min)
  - Vasopressin (units/min)
- Biochemistry
  - Glucose (mmol/L)
  - Highest Lactate in last 48 hrs (mmol/L)
  - Liver function tests Aspartate transaminase/Alanine aminotransferase (U/L)
- Central laboratory specimens
  - Mandatory research blood sample (yes/no)
  - Biobank sample 1 (yes/no)
- SOFA6 score (from day 0 to 14)White Cell Count (x 109/L)
- Delirium (yes/no/sedation/unknown)
- C-reactive protein (mg/L)
- In/out fluids (from day 0 to 14) :

- Total fluids in (mL)
- Total fluids out (mL)
- Balance (mL)

- Process of care/adherence variables:
  - Duration from consent to randomization (minutes)
  - Duration from hospital admission to randomization (minutes)
  - Duration from ICU admission to randomization (minutes)
  - Duration from randomization to start of landiolol (minutes)

Clinical effectiveness analysis

Primary clinical outcome

The primary outcome is the mean SOFA score over 14 days from entry into the trial and whilst in ICU. The SOFA score consists of 6 items each representing an organ system scored from 0-46. The SOFA score is computed by simply summing the scores from the 6 items. A modified version of the SOFA score will be used in this study which excludes the neurology organ dysfunction due to the difficulty in measuring the Glasgow Coma Scale (GCS) i.e., it will have 5 items instead of 6. Hence the score will range from 0-20 instead of 0-24 where a higher score reflects a higher degree of dysfunction/failure.

Secondary clinical outcomes

Mortality at day 28 and day 90 - The mortality status of participants will be collected at day 28 and day 90 and is defined as death due to any cause.

Length of ICU stay and Hospital stay - For those participants discharged alive from ICU, the length of ICU stay will be defined as the time (days) from the date of ICU admission to the date of ICU discharge. Similarly the length of hospital stay for those discharged alive from hospital will be defined as the time (days) from the date of hospital admission to the date of hospital discharge. In addition to this we will assess the time of randomization to ICU discharge and time of randomization to hospital discharge.

Reduction in dose and duration of vasopressor treatment - The dose and duration of noradrenaline and landiolol will be collected for 14 days. The dose and duration of any vasopressin and inotropes administered will be collected during the first five days. There are a number of different inotropes that can be given therefore each will be assessed separately.

Exploratory mechanistic outcomes

Serial blood samples will be collected from participants and analyzed at the Queen Elizabeth Hospital, Birmingham. In particular, the assays will measure markers of myocardial dysfunction7 and inflammation and will be collected at days 0, 1, 2, 4, 6 and end of noradrenaline treatment.

Clinical analysis methods

The results will be reported in accordance with Consolidation Standards of Reporting Trials (CONSORT) guidelines for randomized controlled trials.[8]

The mean and SD for the primary outcome (mean 14-day SOFA score) will be reported by treatment group. Linear mixed effects regression models will be fitted to estimate the treatment effect and 95% confidence interval having adjusted for age, gender, recruiting site (random effect) and baseline noradrenaline dose. An adjustment for the baseline SOFA score will not be required as it will already be included in the mean 14-day SOFA score. If the primary outcome and errors from the linear mixed effects regression model deviate from the normality assumption, then the 95% confidence interval for the mean difference in the SOFA score will be computed using bootstrapping techniques.

Secondary outcomes will also be reported by treatment group reporting the mean and SD for continuous variables and number and percentage for categorical variables. Continuous secondary outcomes will be analyzed in the same way as the primary outcome and the categorical outcomes will be assessed using mixed effects logistic regression models. The reduction in dose of any administered inotropes during the first 5 days will be analyzed using longitudinal models to estimate the difference in reduction over time between the two treatment groups. All models will be adjusted for the same baseline variables as the primary analysis. For all analyses, the adjusted treatment effect estimates will be presented along with their associated 95% confidence interval. For the mortality outcomes at 28 and 90 days, Kaplan-Meier plots will be presented as a visual representation of the time to death. Non-parametric approaches will be considered if any of the secondary outcomes do not satisfy the normality assumptions.

The mechanistic outcomes will be collected on days 0, 1, 2, 4, 6 and end of noradrenaline therapy (EONT). No further samples will be taken after EONT so later samples may not be available. The analyses of the mechanistic data are dependent on the outcomes from the main trial. For that reason, a separate analysis plan will be created for the analysis of the mechanistic outcomes on completion of the main trial. In addition to the primary and secondary analyses, some methodological work will also be conducted. This work will combine mortality and the average SOFA score over a period of 14 days to get a composite endpoint to capture the full extent of the

treatment effect. This composite endpoint would be analyzed using prioritized generalized pairwise comparison methods, namely the win-ratio method/global rank sum technique[9, 10, 11]. Each patient from the intervention group will be compared to each patient in the standard care group (a total of m x n comparisons where m is the total number of patients in the intervention group and n the number of patients in the standard care group) for the death endpoint and then on the average SOFA score. Based on which patient performs better in each pair, the group they belong to would be declared the 'winner'. This would give us the total number of winners in each group and our test statistic would be based on this. In the case of the win-ratio method, for instance, the statistic would be the number of winners in the intervention group divided by the number of winners in the standard care group. This approach will allow us to infer if the intervention is significantly better than the standard care having taken into account the clinical priority i.e. treating mortality as a more important outcome than having a better SOFA score. This work will be reported separate to the main trial results.

Exploratory subgroup analyses will be undertaken for baseline severity (noradrenaline ≥0.1 mcg/kg/min - 0.3 mcg/kg/min vs. >0.3 mcg/kg/min) and use of beta blockers on ICU admission priori to randomization (Yes/No). This will be done using formal statistical tests for interaction for the primary outcome and mortality outcome using mixed effects linear regression models and mixed effects logistic regression models respectively adjusting for the same baseline variables as the primary analysis.

If there is substantial non-adherence, a complier average causal effect (CACE) analysis[13] will be conducted for the primary outcome and mortality. A structural mean model with the inclusion of an instrumental variable will be fitted to estimate the treatment effect among those who complied with the study drug infusion protocol.

A number of other clinical variables collected during the trial will also be compared between the two groups:

- Mean 14 day total fluids in (mL)
- Mean 14 day total fluids out (mL)
- Mean 14 day fluid balance (mL)
- Mean glucose (mmol/L – collected at day 1, day 2 day 4, day 6 and EONT)
- Mean lactate (mmol/L – collected at day 1, day 2 day 4, day 6 and EONT)

Longitudinal models will be fitted for the above outcomes to estimate the adjusted treatment effect and 95% confidence intervals.

The scale and impact of the COVID-19 pandemic on the trial will also need to be assessed on completion of the trial. Baseline demographic and clinical outcomes for participants recruited pre-COVID will be compared to those recruited during COVID to assess if there has been a shift in the population of participants recruited into the trial. For participants recruited during the pandemic, the baseline demographic and clinical outcomes will be compared for participants testing positive for COVID versus those who did not have COVID. Sensitivity analyses will also be performed on the primary and secondary outcomes excluding participants testing positive for COVID to assess the impact on the main trial results.

Handling of missing data

Every effort will be made to minimize missing baseline and outcome data in the trial. The primary outcome is the mean SOFA score over the first 14 days in ICU. This means that a SOFA score is required for each day a patient is alive in ICU, where the SOFA score on any given day is a sum of five items or components. Due to the nature of the study with daily data collection, it is anticipated there will be item level missingness.

If an item is missing then one of the assumptions that can be made is that the measurement might not have been taken as there was no change. Therefore items will be imputed using a last observation carried forward (LOCF) approach as follows (METHOD 1):

- If an item is missing on one or two consecutive days, or if the item is missing at the end of the ICU admission, then the item(s) will be imputed using the LOCF
- If an item is missing on the first day then the value of the item from day two will be taken. If both of these are missing then the baseline value will be used
- If an item is missing for three or more consecutive days, then the mean value of the last observed and next observed values, if available, will be used to impute the missing scores
- If an item is missing for three or more consecutive days including the last day at ICU admission, then the item(s) will be imputed using the LOCF, and
- If an item is missing for three or more consecutive days including day 0, the items(s) will be imputed using the next observed value.

Any differential effect of treatment on ICU discharge or death could affect the comparison of mean SOFA scores between groups. As a post-hoc sensitivity analysis, two further imputation approaches will be conducted for the primary outcome. In the first instance, the last recorded score will be carried forward for all days after ICU discharge or death, up to day 14 (METHOD 2). In the second instance (METHOD 3), all days alive outside of ICU will be assigned the minimum score (zero) and days dead a maximum score (defined as the maximum score that the patient ever obtained). For these sensitivity analyses, the primary analyses will be re-run using the imputed outcome as the dependent variable to estimate the adjusted treatment effect and 95% confidence intervals.

Safety

The number and percentage of SAE and adverse events (AE) in the trial from randomization to 90 day follow-up will be summarized by treatment group. In particular, the following safety outcomes will be recorded:

- Episode of bradycardia (heart rate <50 bpm)
- Bradycardia with hemodynamic compromise requiring significant intervention
- Hypotension requiring significant intervention (not including temporarily stopping the infusion)
- Heart block
- Arrhythmia
- Arrhythmia with hemodynamic compromise requiring intervention

The total number of participants experiencing AE's and SAE's will be compared between the treatment groups using either the chi-squared test or Fisher's exact test for which the P value will be reported. Moreover, the number and percentage will be used to summarize the event type, severity assessment, expectedness and relatedness to intervention by treatment group.

Statistical software

The statistical analyses will be conducted in Stata SE version 16.0.

Authors' contributions

DM and AH drafted the statistical analysis plan and the manuscript which was revised by RL and TW. RL will oversee the final statistical analyses. TW, DY, GDP, DFM, MS, JFB and ACG helped develop the trial design and protocol. RL, DM and SG carried out the power calculations and helped develop the trial design and protocol. JL, MS and TW designed the mechanistic outcomes. TW is the chief investigator and takes responsibility for all aspects of trial design, the protocol and trial conduct. ES oversees the day-to-day running of the trial including protocol compliance. All authors read and approved the final manuscript.

References

1. Fleischmann C, Scherag A, Adhikari NK, et al. Assessment of Global Incidence and Mortality of Hospital-treated Sepsis. Current Estimates and Limitations. *Am J Respir Crit Care Med.* 2016;193(3):259-272.
2. Vincent J-L, Jones G, David S, Olariu E, et al. Frequency and mortality of septic shock in Europe and North America: a systematic review and meta-analysis. *Critical Care.* 2019;23(1):196.
3. Morelli A, Ertmer C, Westphal M, et al. Effect of Heart Rate Control With Esmolol on Hemodynamic and Clinical Outcomes in Patients With Septic Shock: A Randomized Clinical Trial. *JAMA.* 2013;310(16):1683-1691.
4. Gamble C, Krishan A, Stocken D, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *Jama.* 2017;318(23):2337-2343.
5. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomized trials. 2010;340:c332.
6. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med.* 1996;22(7):707-710.
7. Charpentier J, Luyt CE, Fulla Y, et al. Brain natriuretic peptide: A marker of myocardial dysfunction and prognosis during severe sepsis. *Crit Care Med.* 2004;32(3):660-665.
8. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *Jama.* 1996;276(8):637-639.
9. Pocock SJ, Ariti CA, Collier TJ, et al. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *European Heart Journal.* 2011;33(2):176-182.
10. Felker GM, Maisel AS. A global rank end point for clinical trials in acute heart failure. *Circ Heart Fail.* 2010;3(5):643-646.
11. Armstrong PW, Westerhout CM, Van de Werf F, et al. Refining clinical trial composite outcomes: an application to the Assessment of the Safety and Efficacy of a New Thrombolytic-3 (ASSENT-3) trial. *Am Heart J.* 2011;161(5):848-854.
12. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Statistics in medicine.* 2014;33(13):2297-2340.