An international randomised controlled trial of plasma exchange and glucocorticoid dosing in the treatment of anti-neutrophil cytoplasm antibody associated vasculitis: PEXIVAS Trial

Trial Registration: ISRCTN07757494

Statistical Analysis Plan

SAP Version Number
V2.0

Author:	Natalie Ives	Role:	Senior Statistician	Affiliation:	BCTU University of Birmingham
Signature:		Date:			
Chief Investigator:	Dr David Jayne	Role:	Chief Investigator	Affiliation:	Cambridge University Hospitals NHS Foundation Trust
Signature:		Date:			
This Statistical Analysis Plan has been approved by:					
Approver:	Natalie Ives	Role:	Senior Statistician	Affiliation:	BCTU University of Birmingham
Signature:		Date:			

SAP version number	Date Approved	Protocol version number†	Section number changed	Description of and reason for change	Blind Reviewer	
1.0	8th Sept 2011	1.0	N/A	N/A - First release	N/A	
2.0	27th April 2018	3.1		Review of SAP prior to final analysis, decision made to transfer to new BCTU template which is a more comprehensive document. Text from previous SAP copied over to appropriate sections in new template. New sections in template (essentially sections 1-4), uses template text and/or text taken directly from SAP v1.0 or the latest protocol.	Name:	Pollyanna Hardy
					Signature:	
					Date:	
			4.12	Follow-up of patients was changed from minimum of two years to minimum of one year, so timing of the final analysis was changed from 2 years to 1 year in line with when the last patient completed trial follow-up.		
			5.3	Primary analysis populations clarified. The non-inferiority analysis for the GC comparison will be based on the GC per-protocol analysis population (not the ITT population).		
			5.4	The per-protocol population was incorrectly described as an as-treated analysis population. Definition of adherence to PLEX tightened up; receive at least one PLEX within 14 days of randomisation (rather than 28 days). GC adherence calculation: Following advice from the clinical members of the TMG, GC doses given for disease relapse or GC treatment given intravenously should not be included in the GC adherence calculations.		
			8.1	The primary analysis will include covariate adjustment		

				including the two treatment parameters and the minimisation variables. Since the SAP was written, there has been a change in practice with the standard now for analyses to be adjusted for the minimisation variables included in the randomisation process.		
			8.3	Text in SAP v1.0 on handling missing data for the secondary endpoints based on continuous data deleted. These outcomes are being analysed using repeated measures methods which can accommodate missing data.		
			8.5	Clarified that the non-inferiority analysis for the GC comparison will be based on the GC per-protocol analysis population. The primary analysis for the non-inferiority comparison will use a binomial model to calculate the risk difference and 90% Confidence Interval in order to assess the 11% non-inferiority margin. A Cox proportional hazards model will also be fitted to obtain the Hazard Ratio and Confidence Interval. The time to event analysis will be based on a Cox proportional hazards model rather than logrank, as adjusting for the minimisation variables so need to use regression methods.		
			8.6	Analysis methods changed to reflect the use of regression models to adjust for the minimisation variables in the analysis. Serious infections outcome missing in error in SAP v1.0.		
			8.7	Analyses will adjust for the minimisation variables in the models. Sensitivity analyses for CDI censoring for death removed as tertiary outcome. Exploratory analyses restricted to primary outcome (time to event). A third exploratory analysis based on ANCA status added.		

			8.8 8.9	Analysis of time to first SAE removed as not of interest. Subgroup analyses being undertaken for the primary (time to event) outcome only.		

† This SAP was written based on information contained in the trial protocol version as listed here.

Abbreviations & Definitions	
Abbreviation / Acronym	**Meaning**
AAV	ANCA-associated systemic vasculitis
ANCA	Anti-neutrophil cytoplasm antibody
BCTU	Birmingham Clinical Trials Unit
BVAS	Birmingham Vasculitis Activity Score
CONSORT	Consolidated Standards of Reporting Trials
CDA	Combined Damage Assessment Index
DMC	Data Monitoring Committee
eGFR	estimated Glomerular Filtration Rate
EQ-5D	EuroQoL 5D Quality of Life questionnaire
ESRD	End-stage renal disease
GC	Glucocorticoids
ISRCTN	International Standard Randomised Controlled Trial Number
IRR	Incidence rate ratio
ITT	Intention to Treat
PLEX	Plasma Exchange
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SF-36	Short Form-36
TMG	Trial Management Group
Term	**Definition**
End-stage renal disease	The requirement for a renal replacement therapy (hemodialysis or peritoneal dialysis) for at least 12 consecutive weeks or the receipt of a renal transplant.
International Standard Randomised Controlled Trial Number	A clinical trial registry
Protocol	Document that details the rationale, objectives, design, methodology and statistical considerations of the study
Randomisation	The process of assigning trial subjects to treatment or control groups using an element of chance to determine the assignments in order to reduce bias.
Statistical Analysis Plan	Pre-specified statistical methodology documented for the trial, either in the protocol or in a separate document.

TABLE OF CONTENTS

1. Introduction

This document is the Statistical Analysis Plan (SAP) for the PEXIVAS trial, and should be read in conjunction with the current trial protocol. This SAP details the proposed analyses and presentation of the data for the main paper(s) reporting the results for the PEXIVAS trial.

The results reported in these papers will follow the strategy set out here. Subsequent analyses, which will include both specifically-planned secondary analyses and other analyses of a more exploratory nature will not be bound by this strategy, though they are expected to follow the broad principles laid down here. The principles are not intended to curtail exploratory analysis (e.g. to decide cut-points for categorisation of continuous variables), nor to prohibit accepted practices (e.g. transformation of data prior to analysis), but they are intended to establish rules that will be followed, as closely as possible, when analysing and reporting data.

Any deviations from this SAP will be described and justified in the final report or publication of the trial (using a table as shown in Appendix A). The analysis will be carried out by an appropriately qualified statistician, who should ensure integrity of the data during their data cleaning processes.

2. Background and rationale

The background and rationale for the trial are outlined in detail in the protocol. In brief, anti-neutrophil cytoplasm antibody (ANCA)-associated systemic vasculitis (AAV) has a universally poor prognosis with mortality approaching 100% within 5 years. The introduction of treatment regimens based on cyclophosphamide and glucocorticoids (GC) have transformed AVV from a rapidly fatal disease to one of chronic morbidity and reduced survival often preceded by end-stage renal disease (ESRD).

Plasma exchange (PLEX), a method of rapidly removing potentially pathogenic ANCA and other mediators of inflammation and coagulation, has shown promise as an adjunctive therapy in AAV to improve early disease control and improve rates of renal recovery in severe disease. Glucocorticoids are a standard of care in the treatment of AVV. High doses of GC early in disease although undeniably reduce disease activity due to their anti-inflammatory and immunosuppressive properties also increase the risk of infection.

There is a need for therapies with reduced toxicity while improving disease control.

3. Trial objectives

The primary objectives are to:

1. determine the efficacy of PLEX in addition to immunosuppressive therapy and GC with respect to death and ESRD.
2. determine whether a reduced dose GC regimen is non-inferior to a standard dose regimen with respect to death and ESRD.

Secondary objectives are limited to those of direct relevance to the assessment of efficacy and safety of the treatments. For both i) PLEX in addition to immunosuppressive therapy and GC compared to immunosuppressive therapy and GC alone and ii) reduced dose GC compared to standard dose GC, the secondary objectives are to:

1. determine the effect on disease activity
2. determine the effect on death
3. determine the effect on ESRD
4. determine safety
5. determine effects on serious infections
6. determine effects on health related quality of life

Exploratory objectives for both i) PLEX in addition to immunosuppressive therapy and GC compared to immunosuppressive therapy and GC alone and ii) reduced dose GC compared to standard dose GC are to:

1. determine the cost-effectiveness of each treatment compared to the standard of care
2. determine the effects on measures of disease related damage
3. determine the effects on long-term renal function

4. Trial methods

4.1. Trial design

PEXIVAS is an international multi-centre, prospective, open-label, 2x2 factorial design, phase III randomised controlled trial in severe AAV. Each participant will be follow-up until the study close with a minimum duration of follow-up of approximately 1 year. See Appendix B for trial schema.

4.2. Trial interventions

Participants will be randomised to either adjunctive PLEX or no PLEX and randomised to either reduced dose GC or standard dose GC. The four treatment groups are therefore:

PLEX in additional to standard immunosuppressive therapy and GC;

Standard immunosuppressive therapy and GC (no PLEX group);

Reduced dose GC taper; and

Standard dose GC taper.

PLEX will consist of 7 exchanges within 14 days of randomisation, of at least 60ml/kg (based on actual body weight) per session using albumin as a replacement solution.

Oral GC will consist of non-enteric coated prednisone or prednisolone at equivalent mg to mg doses. Dosing will depend on the participant's weight at randomisation with three possible weight categories. All oral GC will be given as a single daily dose. The dosing for the reduced and standard dose GC groups is given in the Table below (this is Table 1 in the protocol).

Week	Standard			Reduced-dose		
	<50 kg	50-75 kg	>75 kg	<50 kg	50-75 kg	>75 kg
	pulse	pulse	pulse	pulse	pulse	pulse
1	50	60	75	50	60	75
2	50	60	75	25	30	40
3-4	40	50	60	20	25	30
5-6	30	40	50	15	20	25
7-8	25	30	40	12.5	15	20
9-10	20	25	30	10	12.5	15
11-12	15	20	25	7.5	10	12.5
13-14	12.5	15	20	6	7.5	10
15-16	10	10	15	5	5	7.5
17-18	10	10	15	5	5	7.5
19-20	7.5	7.5	10	5	5	5
21-22	7.5	7.5	7.5	5	5	5
23-52	5	5	5	5	5	5
>52	Investigators' Local Practice			Investigators' Local Practice		

In summary, all participants will receive either 50, 60 or 75mg/day (based on weight at randomisation) of oral GC for the first 7 days.

- Participants in the reduced dose group will continue at 25, 30 or 40mg/day for the next 7 days and then taper to between 6 and 10mg/day by 3 months and 5mg/day by 6 months.
- Participants in the standard dose group will continue at 50, 60 or 75mg/day for the next 7 days and then taper to between 12.5 and 20mg/day by 3 months and 5mg/day at 6 months.

All participants will receive 5mg/day from 6 months to 12 months after randomisation.

The reduced dose regimen will expose participants to approximately 50% of the standard oral

dose over the first 3 months and 53% over the first 6 months of treatment.

4.3. Primary outcome measure

The primary outcome is a composite of i) all-cause mortality or ii) ESRD.

ESRD is defined as the requirement for a renal replacement therapy (hemodialysis or peritoneal dialysis) for at least 12 consecutive weeks or the receipt of a renal transplant.

4.4. Secondary outcome measures

Secondary outcomes are as follows:

1. Sustained remission (defined as remission that is obtained within 6 months of randomisation and lasts without a relapse until at least 12 months after randomisation);
2. All-cause mortality;
3. ESRD;
4. Serious adverse events (SAEs) defined as any medical occurrence that results in permanent disability, hospitalisation or prolongation of a hospitalisation, is life threatening or results in death;
5. Serious infections defined as an infectious syndrome that requires intravenous antibiotics or hospitalisation for treatment;
6. Medical Outcomes Survey Short Form-36 (SF-36) Physical Composite Score and Mental Composite Score;
7. EuroQoL EQ-5D Index Score (3 level version).

The SF-36 consists of eight scaled scores, and a Physical Composite Score and Mental Composite Score. Each scale is directly transformed into a 0-100 scale. The lower the score the more disability i.e., a score of zero is equivalent to maximum disability and a score of 100 is equivalent to no disability (see Section 8.4).

The EQ-5D has five dimensions which are combined into a 5-digit number that describes the patient's health state. This health state is transformed into a score that ranges from -0.59 (worse health) to 1 (best health). The EQ-5D also contains a 100-point visual analogue scale, where 0=worst health you can imagine and 100=best health you can imagine (see Section 8.4).

4.5. Tertiary outcome measures

Tertiary outcomes are as follows:

1. Cost-effectiveness ratios (not part of this SAP);
2. Combined Damage Assessment Index (CDA); the physician global assessment of

damage;

3. Estimated Glomerular Filtration Rate (eGFR; Modification of Diet in Renal Disease four variable formula).

4.6. Timing of outcome assessments

The schedule of trial procedures are given in Appendix C (see also Table 4 in the protocol).

4.7. Randomisation

Participants will be randomised to the treatments in a one-to-one ratio to either PLEX or no PLEX and reduced dose GC or standard dose GC.

Randomisation will be performed by a central randomisation facility utilising a computerised minimisation algorithm. The algorithm will not be made available to investigators. Allocation will follow a minimisation scheme based on the following prognostic factors in AAV:

- Severity of renal disease at presentation (requiring dialysis or creatinine ≥500µmol/L (5.6mg/dL) vs. <500µmol/L);
- Age (<60 vs. ≥60 years old);
- ANCA subtype (PR3-ANCA vs. MPO-ANCA);
- Severity of lung hemorrhage (no hemorrhage vs. lung hemorrhage with a blood oxygen saturation of >85% on room air vs. lung hemorrhage with blood oxygen saturation of ≤85% on room air or ventilated);
- Planned induction immunosuppressive therapy to be used (oral CYC vs. intravenous CYC vs. rituximab).

4.8. Sample size

The sample size for this trial is event driven in order to detect a hazard ratio of 0.64 (PLEX vs. no PLEX) with 80% power and a 2-sided alpha of 0.05. Protocol versions 1.0 and 2.0 estimated a required sample size of 500 predicting 164 events over the study period equivalent to a 12% absolute risk reduction of the primary endpoint at 5 years (44% in the no PLEX group vs. 32% in the PLEX group; overall 38%). This sample size estimate assumed a 5 year median time to ESRD or death on the basis of previous extended follow-up studies in randomised trials of AAV of a similar severity to those targets in this study.

Review of the PEXIVAS events rates in 2014 indicated a 2 year event rate of 24% and predicted overall 5 year event rate of 30-35%. Improvements in death and ESRD have been recently reported in registry studies. In order to obtain the required number of events, the sample size needed to be increased to 675-725 participants allowing for a 10% loss to follow-up or cross over between treatment groups. The trial planned to enroll 700 participants in order to observe at least 160 events. These calculations assume no significant interaction

between the two treatment factors. Although the absolute risk appears larger than is often considered clinically important, the expensive and invasive nature of the primary intervention, PLEX warrants a relatively large effect size. Additionally, this effect size is close to the estimated effect of PLEX in the meta-analysis of prior studies (80% power to detect a relative risk reduction of 27% with our sample size compared to a relative risk reduction of 20% in the meta-analysis).

While this effect size appears reasonable to detect for PLEX, it is unlikely a reduction in GC will result in a 12% absolute risk reduction of death or dialysis. However, we expect approximately 25% of participants to experience a severe infection based on prior studies. A sample size of 700 participants will allow 80% power to detect at least a 10% absolute risk reduction in severe infections (relative risk reduction of severe infection by 40%), a finding of clinical importance.

In terms of the non-inferiority hypothesis, a sample size of 700 participants would allow >80% power to ensure that the reduced dose GC regimen results in an increase in ESRD or death by no more than 11% (one-sided alpha of 0.05).

4.9. Framework

The objective of the trial is to test the superiority of PLEX vs. no PLEX and the non-inferiority of reduced dose GC vs. standard dose GC for the primary outcome of all-cause mortality and/or ESRD.

The comparison of PLEX to no PLEX is a superiority hypothesis expressed as the PLEX group relative to the no PLEX group. The null hypothesis is that there is no difference in the composite outcome of all-cause mortality and ESRD between the treatment groups. The alternative hypothesis is that there is a difference between the groups.

The comparison of reduced dose GC to standard dose GC is a non-inferiority hypothesis with a non-inferiority margin of an 11% absolute risk increase expressed as the reduced dose GC group relative to the standard dose GC group. An intention to treat analysis can increase the risk of falsely claiming non-inferiority,[1] therefore we will conduct the primary analysis using only participants adherent to the assigned GC regimen (the per-protocol population, see section 5.4), with a sensitivity analysis utilising all participants (the intention to treat (ITT) population, see section 5.4). The absolute risks in each group will be calculated using the complete follow-up data and the 90% confidence interval will be computed to correspond to a one-sided alpha of 0.05. If the 90% confidence interval excludes an 11% increase in the composite primary outcome, the inference will be non-inferiority.

4.10. Interim analyses and stopping guidance

A separate Data Monitoring Committee (DMC) reporting template will be drafted and agreed by the DMC including an agreement on which outcomes will be reported at interim analyses.

Interim analyses of efficacy and safety are planned annually. The Haybittle-Peto approach will be used whereby all interim analyses use a difference of 3 standard errors (approximately p=0.002) as a stopping guideline. Efficacy and safety data will be reviewed by the DMC on an annual basis or more frequently if required by the DMC or Trial Management Group.

4.11.Internal Pilot Progression Rules

Not applicable.

4.12. Timing of final analysis

The final analysis of all the trial data for the main publication purposes will occur approximately one year after the final participant has been entered into the trial and the corresponding outcome data has been entered onto the trial database and validated as being ready for analysis. This is provided the trial has not stopped recruitment early for any reason (e.g. DMC advice or funding body request).

4.13. Timing of other analyses

Additional analyses of the trial data will be conducted in the future but are beyond the scope of this SAP which is documenting the analyses that will be undertaken by BCTU for the end of trial analyses.

4.14.Trial comparisons

PEXIVAS has a 2x2 factorial design, so has two main effect comparisons. The primary comparison groups will be composed of:

- those randomised to PLEX and those randomised to no PLEX (comparison 1)
- those randomised to a reduced dose GC regimen and those randomised to a standard dose GC regimen (comparison 2).

5. Statistical Principles

5.1. Confidence intervals and p-values

Unless otherwise specified, all estimates of differences between groups will be presented with two-sided 95% confidence intervals. P-values will be reported from two-sided tests at the 5%

significance level.

5.2. Adjustments for multiplicity

No correction for multiple testing will be made.

5.3. Analysis populations

All primary analyses (primary, secondary and tertiary outcomes) will be by ITT, except the primary analysis of the primary outcome for the GC dose comparison which will be based on the GC per-protocol analysis population (as this is a non-inferiority analysis).

The ITT analysis population will include all randomised participants whether or not they received the allocated treatment, and whether or not they were withdrawn/ineligible, and participants will be analysed in the treatment group to which they were randomised. This is to avoid any potential bias in the analysis.

An ITT analysis can increase the risk of falsely claiming non-inferiority,[1] therefore for the assessment of whether a reduced dose GC regimen is non-inferior to a standard dose GC regimen with respect to the primary outcome (only), a per-protocol analysis will be the primary analysis (see section 5.4). For the secondary and tertiary outcomes, the analysis will be based on the ITT analysis population.

There is a risk of cross-over of patients randomised to not receive PLEX to the PLEX arm, and vice-versa, although this practice will be discouraged. Similarly there is a risk that patients may receive a dose of GC appreciably different to that which they were allocated. The primary ITT and per-protocol analyses will deal with these cross-overs in a conservative manner (bias to the null). To explore the potential that these cross-overs may reduce the true magnitude of treatment effects, sensitivity analyses will be performed. These analyses will include per-protocol and ITT analyses (for the PLEX and GC comparisons respectively); the per-protocol populations are defined in section 5.4. Due to the inherent potential for bias, these sensitivity analyses will not, irrespective of any differences to the primary analyses, supplant the planned primary analyses.

5.4. Definition of adherence

<u>Adherence to PLEX</u>

Participants randomised to PLEX will be regarded as having received PLEX if they received at least one complete exchange within 14 days of randomisation.

The PLEX vs. no PLEX per-protocol analysis population will consist of:

- PLEX group: Participants randomised to PLEX who receive at least one complete exchange within 14 days of randomisation or who died within 14 days of randomisation.

- No PLEX group: Participants randomised to no PLEX who received no PLEX.

All per-protocol analyses comparing the PLEX vs. no PLEX groups will be based on this population irrespective of adherence to the GC.

Adherence to GC

Participants will be regarded as receiving the reduced dose of GC if they receive no more than 130% of the cumulative oral dose of the reduced dose regimen in the first 6 months of therapy.

Participants will be regarded as receiving the standard dose of GC if they receive at least 70% of the cumulative oral dose of the standard regimen in the first 6 months of therapy.

For the adherence calculation, it was decided by the clinical members of the Trial Management Group that any GC doses administered due to suspected disease activity (i.e. relapse) or any GC treatment given to participants that was administered intravenously were not considered part of the trial GC regimen, and so did not contribute to the assessment of adherence.

The GC dose per-protocol analysis population will consist of:

- Participants randomised to reduced dose GC who receive ≤130% of the cumulative oral dose of the reduced dose regimen in the first 6 months of therapy.
- Participants randomised to standard dose GC who receive ≥70% of the cumulative oral dose of the standard regimen in the first 6 months of therapy.

All per-protocol analyses comparing the reduced vs. standard GC groups will be based on this population irrespective of adherence to PLEX.

5.5. Handling protocol deviations and violations

A protocol deviation/violation is defined as a failure to adhere to the protocol such as errors in applying the inclusion/exclusion criteria, the incorrect intervention being given, incorrect data being collected or measured, follow-up visits outside the visit window or missed follow-up visits. We will apply a strict definition of the ITT principle and will include all participants as per the ITT population described in section 5.4 in the analysis in some form regardless of deviation from the protocol.[2] This includes participants who were randomised but later found to violate the inclusion or exclusion criteria. This does not include those participants who have specifically withdrawn consent for the use of their data in the first instance.

5.6. Unblinding

Not applicable.

6. Trial population

6.1. Recruitment

A flow diagram (as recommended by CONSORT[3]) will be produced to describe the participant flow through each stage of the trial. This will include information on the number (with reasons) of losses to follow-up (drop-outs and withdrawals) over the course of the trial.

6.2. Baseline characteristics

The baseline characteristics of the trial populations for the two trial comparisons (PLEX vs. no PLEX and reduced vs. standard dose GC) will be summarised. Categorical data will be summarised by number of participants, counts and percentages. Continuous data will be summarised by the number of participants, mean and standard deviation if deemed to be normally distributed or number of participants, median and interquartile range if data appear skewed, and ranges if appropriate. Tests of statistical significance will not be undertaken, nor confidence intervals presented.[4]

7. Intervention(s)

7.1. Description of the intervention(s)

Further information on the PLEX intervention and GC doses will be provided. This will include a tabulation of the number of PLEX exchanges received in participants in the PLEX vs. no PLEX comparison. Summary measures of dose of GC received will be reported for the reduced vs. standard GC dose comparison.

7.2. Adherence to treatment allocation

A cross-tabulation of allocated treatment by the adherence categories stated in section 5.4 will be produced (counts and percentages).

8. Analysis methods

PEXIVAS has a 2x2 factorial design, so has two main effect comparisons. The primary comparison groups will be composed of:

- those randomised to PLEX and those randomised treated to no PLEX (comparison 1)
- those randomised to the reduced-dose glucocorticoid regimen and those randomised to the standard-dose glucocorticoid regimen (comparison 2).

Unless otherwise specified, the two comparisons will be analysed by the same methods, and all analyses described below will be carried out twice, once for each comparison. It is expected *a priori* that no interaction between the treatments will be identified, but if one is found, this will

be considered a chance finding. So whilst an interaction test will be carried out, it shall be interpreted with caution. To avoid the increased risk of a false positive result associated with multiple testing, an interaction test will only be carried out for the primary outcome.

Treatment groups will be compared using generalised estimating equations, or a similar method, to adjust for all covariates as specified in section 8.1, where possible.

8.1. Covariate adjustment

In the first instance, treatment effects between groups for all outcomes will be adjusted for both treatment group parameters (i.e. PLEX vs. no PLEX and reduced vs. standard dose GC) and the minimisation parameters listed in section 4.7.[5] Categorised continuous variables (e.g. age) will be treated as continuous variables in this adjustment. Other covariate adjustment will be for baseline values for parameters where available (e.g. the repeated measures analysis of the SF-36 and EQ-5D will also include the relevant baseline score as a covariate).

If covariate adjustment is not possible (e.g. the model does not converge), the minimisation variables will be removed first to produce a partially adjusted model with just the two treatment group parameters included as covariates. If the model still does not converge, then unadjusted estimates will be produced, and it will be made clear in the final report why this occurred (e.g. not possible due to low event rate/lack of model convergence).

8.2. Distributional assumptions and outlying responses

Distributional assumptions (e.g. proportional hazards assumption) will be assessed visually prior to analysis; although in the first instance the proposed primary method of estimation in this analysis plan will be followed.

8.3. Handling missing data

In the first instance, analysis will be completed on received data only with every effort made to follow-up participants even after protocol violation to minimise any potential for bias.

For the primary analyses, no attempt will be made to impute missing data. The primary outcome is time-to-event data and thus, the standard methods of analysis naturally allow for early drop-out.

8.4. Data manipulations

The Trial Statistician will derive all responses from the raw data recorded in the database.

The primary outcome is a composite of i) all-cause mortality or ii) ESRD. This will be analysed as both a time to event outcome and a binary (yes/no) outcome.

The time to event primary outcome is the time between the date of randomisation and date of death or date of ESRD, whichever occurs first. For participants who are alive without ESRD, the date of withdrawal, date lost to follow-up or date of last clinic follow-up assessment will be used for censoring. Participants who died on the same day as randomisation will be counted as having the event on day 1. Participants who have either withdrawn or are lost to follow-up on the same day as randomisation will be censored at day 1.

The binary primary outcome will use all trial data over the whole trial follow-up period, with participants who die or reach ESRD being counted as having an event (yes).

All-cause mortality is defined as time from date of randomisation to date of death. For participants who are alive, the date of withdrawal, date lost to follow-up or date of last clinic follow-up assessment will be used for censoring. Participants who died on the same day as randomisation will be counted as having the event on day 1. Participants who have either withdrawn or are lost to follow-up on the same day as randomisation will be censored at day 1.

ESRD is defined as time for randomisation to date of ESRD. For participants who do not reach ESRD, the date of death, date of withdrawal, date lost to follow-up or date of last clinic follow-up assessment will be used for censoring. Participants who reach ESRD on the same day as the randomisation will be counted as having the event on day 1. Participants who have either died, withdrawn or are lost to follow-up on the same day as randomisation will be censored at day 1.

Serious infections are reported on both the infection and SAE forms. The data on infections from these forms will be collated to determine whether participant's experienced serious infections.

The SF-36 consists of eight scaled scores, which are the weighted sums of the questions in their section, and a Physical Composite Score and Mental Composite Score. Each scale is directly transformed into a 0-100 scale on the assumption that each question carries equal weight. The lower the score the more disability. The higher the score the less disability i.e., a score of zero is equivalent to maximum disability and a score of 100 is equivalent to no disability. The eight sections are:

- Vitality
- Physical functioning
- Bodily pain
- General health perceptions
- Physical role functioning

- Emotional role functioning
- Social role functioning
- Mental health

The EQ-5D has five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 3 levels: no problems, some problems, and extreme problems. The digits from the five dimensions can be combined into a 5-digit number that describes the patient's health state, which is then transformed into a score that ranges from -0.59 (worse health) to 1 (best health). The EQ-5D also contains a 100-point visual analogue scale, where 0=worst health you can imagine and 100=best health you can imagine.

The CDA index is for recording organ damage that has occurred in patients since the onset of vasculitis. Damage is irreversible, and only rarely should a scored item not be carried forward. Since CDA is a tertiary outcome, all items will be carried forward regardless of how that item is answered on subsequent forms. The CDA has a maximum score of 116.

8.5. Analysis methods – primary outcome

The primary outcome is a composite of i) all-cause mortality or ii) ESRD.

The comparison of PLEX vs. no PLEX is based on the ITT analysis population using a time to event analysis (time from randomisation to death or ESRD). The primary outcome will be compared between treatment arms using survival analysis methods. Kaplan-Meier survival curves will be constructed for visual presentation of time-to-event comparisons. A Cox proportional hazards model will be fitted to obtain an adjusted hazard ratio and 95% confidence interval. The analyses will be adjusted using the covariate adjustment as described in section 8.1. In order to assess the impact of including covariates in the model, the partially adjusted model (with the two treatment group parameters) will also be presented.

The comparison of reduced vs. standard dose GC is a non-inferiority hypothesis with a non-inferiority margin of an 11% absolute risk increase. This analysis will be based on the GC per-protocol analysis population. The absolute risks in each group of participants reaching the primary outcome using the binary outcome of death or ESRD will be calculated using the complete follow-up data. A binomial model will be fitted to obtain the adjusted risk difference and 90% confidence interval (adjusting for the covariates listed in section 8.1). If the 90% confidence interval excludes an 11% increase in the primary outcome, the inference will be non-inferiority. In addition to the absolute risk difference, we will calculate the adjusted hazard ratio and 95% confidence interval from the Cox proportional hazards model using the ITT analysis population. In order to assess the impact of including covariates in the model, the partially adjusted model (with the two treatment group parameters) will also be presented.

8.6. Analysis methods – secondary outcomes

For both trial comparisons, all analyses of secondary outcome measures will be based on the ITT analysis population.

Sustained Remission

Disease activity will be analysed in terms of sustained remissions. Participants will have obtained a sustained remission if they achieve a Birmingham Vasculitis Activity Score (BVAS)/WG of zero (complete remission) within 26 weeks of randomisation and maintain a BVAS/WG of zero without evidence of relapse from complete remission until at least 52 weeks after randomisation. The number and percentage of participants that achieve a sustained remission will be calculated for each treatment arm and adjusted relative risks and 95% confidence intervals will be estimated using a log binomial regression model, or using a log Poisson regression model with a robust variance estimator if the binomial model fails to converge.[6]

ESRD and Death

ESRD and death will be analysed as separate endpoints. Kaplan-Meier curves will be constructed for visual presentation and a Cox proportional hazards model will be fitted to obtain an adjusted hazard ratio and 95% confidence interval.

Serious Infections

The rate of serious infections will be assessed both for the first year of the trial and at trial end. The number of serious infections a participant experienced will be analysed using a Poisson regression or negative binomial regression, with an offset for the length of time the participant was in the trial included in the model, to obtain an adjusted incidence rate ratio (IRR) and 95% confidence interval.

Quality of Life Measures

Health-related quality of life will be assessed using the SF-36 Physical Composite and Mental Composite Scores, together with the individual SF-36 domain scores. Longitudinal plots of the mean scores and mean changes from baseline over time by treatment group will be produced for visual inspection of the data. Data will be analysed using mixed effect repeated measures models with the two treatment group parameters, minimisation variables and the baseline questionnaire score included in the model as covariates. Time will be included as a continuous variable in the model. In the initial model, a treatment by time cross-term will be included in the model. If this is not significant, it will be considered that the treatment effect is constant over time, and models without the treatment by time cross-term will be fitted.

Separate models will be fitted for each of the SF-36 composite and domain scores. Due to the increased risk of false positive results with multiple testing, the composite scores will dominate the interpretation, and analysis of the individual domain scores will be considered exploratory.

For the EQ-5D, separate models for the Index and Visual Analogue Score will be fitted.

8.7. Analysis methods – exploratory outcomes and analyses

Tertiary Outcomes

For both trial comparisons, all tertiary analyses will be based on the ITT analysis population.

Combined Damage Assessment Index

Damage will be compared between groups using serial CDA index scores as the outcome in a mixed effect repeated measures models with the two treatment group parameters, minimisation variables and the baseline CDA value included in the model as covariates. Time will be included a continuous variable in the model. In the initial model, a treatment by time cross-term will be included in the model. If this is not significant, it will be considered that the treatment effect is constant over time, and models without the treatment by time cross-term will be fitted.

Renal Function

Renal function will be compared between groups using serial eGFR as the outcome in a mixed effect repeated measures model with the two treatment group parameters, minimisation variables and baseline eGFR value included in the model as covariates. Time will be included as a continuous variable in the model. In the initial model, a treatment by time cross-term will be included in the model. If this is not significant, it will be considered that the treatment effect is constant over time, and models without the treatment by time cross-term will be fitted. Participants who reach ESRD will be considered to have an eGFR of zero for all subsequent time points. Participants who die will be censored at that point, unless they die whilst on dialysis in which case they will remain in the analysis with eGFR of zero. Longitudinal plots of mean values and mean changes from baseline over time by treatment group will be produced for visual presentation of the data.

Exploratory Analyses

An exploratory analysis will be carried out to investigate for the primary outcome (time to event) whether the method of PLEX treatment affects the results. This will be a non-randomised comparison and thus subject to selection bias. Although relevant covariates will be included in the analysis in an attempt to mitigate any bias, these results will be treated as hypothesis generating only and considered of tertiary importance in the trial.

Due to the potential that the investigational treatments may largely affect early mortality and renal function, analyses of the primary outcome (the time to event outcome of time to death or ESRD) will also be performed after censoring data at 12 months follow-up.

As a putative mediator of disease, it is important to understand if ANCA levels are affected by plasma exchange or steroid dosing. Therefore, the proportion of participants who are ANCA

negative at 2 weeks and at 4 weeks will be calculated for each treatment comparison.

8.8. Safety data

The number and percentage of participants experiencing at least one SAE will be analysed as a categorical binary (yes/no) variable and adjusted relative risks and 95% confidence intervals will be estimated using a log binomial regression model, or using a log Poisson regression model with a robust variance estimator if the binomial model fails to converge.[6] If numbers allow, a more complex model of SAE occurrences will be constructed utilising SAE as a count variable. The number of SAEs a participant experienced will be analysed using a Poisson regression or negative binomial regression, with an offset for the length of time the participant was in the trial included in the model, to obtain an adjusted incidence rate ratio (IRR) and 95% confidence interval.

An analysis of major subgroups of SAEs (e.g. infections, malignancy, cardiovascular complications) will be performed separately, but in an identical manner as the overall SAE analyses.

8.9. Planned subgroup analyses

Several *a priori* subgroup analyses are planned with respect to the primary outcome (time to event) using the ITT analysis populations for both the PLEX vs. no PLEX and reduced dose vs standard dose GC comparisons. The subgroups will be each of the strata included in the randomisation minimisation variables:

- Severity of renal disease at presentation (requiring dialysis or creatinine ≥500µmol/L (5.6mg/dL) vs. <500µmol/L);
- Age (<60 vs. ≥60 years old);
- ANCA binding specificity (PR3-ANCA vs. MPO-ANCA);
- Severity of lung hemorrhage (no hemorrhage vs. hemorrhage with blood oxygen saturation >85% on room air vs. hemorrhage with blood oxygen saturation ≤ 85% on room air or ventilated);
- Planned induction immunosuppression therapy to be used (IV CYC vs. oral CYC vs. rituximab).

The effects of these subgroups will be examined by including the relevant subgroup by treatment interaction term in the Cox proportional hazards model for each separate subgroup analysis to explore whether there is evidence that the treatment effects (PLEX vs. no PLEX or reduced vs. standard dose GC) differ across subgroups. Tests of heterogeneity will be presented along with subgroup specific hazard ratios and 95% confidence intervals. Interpretation of subgroup analysis will be treated with caution (output will be treated as exploratory rather than definitive[7]).

8.10. Sensitivity analyses

Sensitivity analyses will be limited to the primary outcome and will consist of a:

- Per-protocol analysis for the PLEX vs. no PLEX trial comparison using the per-protocol population as described in section 5.4.

9. Analysis of sub-randomisations

Not applicable.

10. Health economic analysis

As indicated in the protocol there will also be an economic analysis. The details of this analysis will be documented separately.

11. Statistical software

Statistical analysis will be undertaken in the following statistical software packages SAS and Stata.

12. Differences to the protocol

SAP v1.0 reflected the text in the protocol v3.1. The changes outlined in the table at the front of this SAP reflect the changes from v1.0 to v2.0, and thus include differences to the protocol. The analysis methods described in this SAP will be followed; where the methods in the latest protocol differ, this SAP will be followed.

13. References

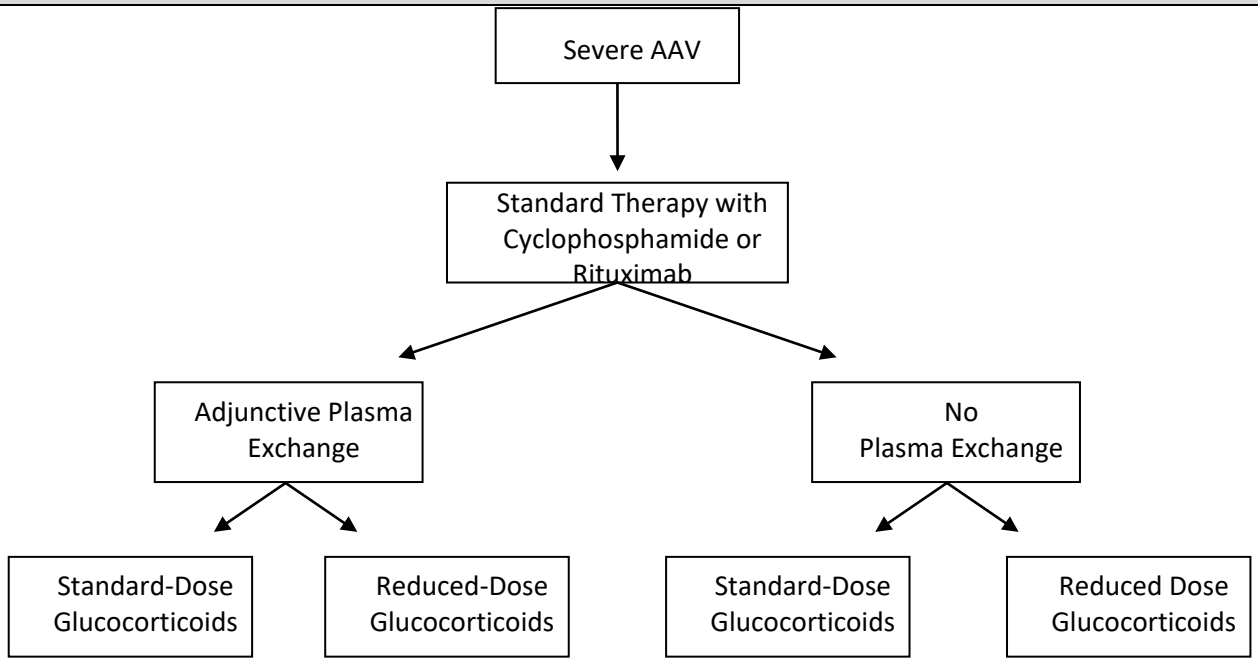
1. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. JAMA. 2006;295(10):1152-60.
2. Gupta SK. Intention-to-treat concept: A review. Perspect Clin Res. 2011;2(3):109-112.
3. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
4. Altman DG, Dore CJ. Randomisation and baseline comparisons in clinical trials. Lancet. 1990;335:149–53.
5. Montgomery AA, Peters TJ, Little P. Design, analysis and presentation of factorial randomised controlled trials. BMC Med Res Methodol. 2003;3:26.
6. Zou G. A modified Poisson regression approach to prospective studies with binary data. Am J Epidemiol. 2004;159(7):702-6.
7. Wand R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Reporting of subgroups analyses in clinical trials. NEJM. 2007;357:2189-94.

Appendix A: Deviations from SAP

This report below follows the statistical analysis plan dated <insert effective date of latest SAP> apart from following:

Section of report not following SAP	Reason
<insert section >	<insert, e.g. exploratory analyses request by TMG>

Appendix B: Trial schema



Appendix C: Schedule of assessments

	Induction of Remission								Maintenance of Remission				
Study Visits	**Screen**	**Baseline**		**WK 2**	**WK 4**	**WK 8**	**WK 12**		**WK 26**	**WK 39**	**WK 52**	**Every 26 WKs until Study Termination**	**Relapse Visit *or* Termination Visit**
Study Drug													
Glucocorticoid Dose	X	X		X	X	X	X		X	X	X	X	X
PLEX Type/Details				X									
Data Forms													
Informed Consent	X												
Eligibility	X												
Randomization		X											
Demographics	X												
Clinical Data	X			X	X	X	X		X	X	X	X	X
Weight		X		X	X	X	X		X	X	X	X	X
Medications	X	X		X	X	X	X		X	X	X	X	X
BVAS/WG	X	X		X	X	X	X		X	X	X	X	X
CDA		X					X		X		X	X	X
SF-36 and EQ5D		X					X		X		X	X	X
Adverse Event Report				X	X	X	X		X	X	X	X	X
Clinical Labs													
ANCA	X			X			X		X	X	X	X	X
Anti-GBM	X												
Creatinine	X	X		X	X	X	X		X	X	X	X	X
Pregnancy Test*	X												

Research Specimens													
DNA**		X											
RNA		X					X				X		
Serum		X		X			X				X		
Plasma		X		X			X				X		
Renal pathology§		X											

Appendix D: Template report

A template report for the final analyses will be provided in a separate document.