A randomised controlled trial of a combined screening /treatment programme to prevent premature failure of renal transplants due to chronic rejection in patients with HLA antibodies (OuTSMART trial)

Statistical Analysis Plan

Version 2.4

Version 2.4 started: 09/02/2021

ISRCTN number: ISRCTN46157828

Trial Statistician: Dominic Stringer

Signature..... Date.....

Chief Investigator: Professor Anthony Dorling

Signature..... Date.....

Trial Steering Committee Chair: Professor Chris Watson

Signature..... Date.....

Purpose and Scope of Statistical Analysis Strategy

This document details the presentation and analysis strategy for the primary papers reporting results from the OUTSMART trial. It is intended that the results reported in these papers will follow the strategy set out herein; subsequent papers of a more exploratory nature will not be bound by this strategy but will be expected to follow the broad principles laid down for the principal paper(s). These principles are not intended to curtail exploratory analysis or to prohibit sensible statistical and reporting practices, but they are intended to establish the strategy that will be followed as closely as possible, when analysing and reporting the trial. Reference was made to the trial protocol (OUTSMART Trial Protocol V14 08_07_2020), ICH [44] guidelines on Statistical Principles (E9) and CONSORT [45] guidelines.

Contents

A) QUANTITATIVE ANALYSIS PLAN

Investigators:

Principal investigator: Anthony Dorling

Trial manager: Leanne Gardner

Trial statisticians: Dominic Stringer, Janet Peacock

Trial health economist: Paul McCrone

Description of the trial

Protocol Publication:

Anthony Dorling, Irene Rebollo Mesa, Rachel Hilton, Janet Peacock, Robert Vaughan, Richard Baker, Brendan Clarke, Raj Thuraisingham, Matthew Buckland, Michael Picton, Susan Martin, Richard Borrows, David Briggs, Robert Horne, Paul McCrone and Caroline Murphy. Can a combined screening /treatment programme prevent premature failure of renal transplants due to chronic rejection in patients with HLA antibodies: Study protocol for the multicentre randomised controlled OuTSMART trial. (Trials. 2014 Jan 21; 15:30. doi: 10.1186/1745-6215-15-30)

Principal research objectives to be addressed

Primary objective;

Compare the time to graft failure in patients with HLA Ab who receive an optimized anti-rejection medication intervention ('treatment'), with that in a control group with HLA Ab who remain on their established immunotherapy and whose clinicians are not aware of their Ab status.

Secondary objectives;

a) Determine the time to graft failure in patients randomized to 'unblinded' HLA Ab screening, compared to a control group randomized to 'blinded' HLA Ab screening.

b) Determine whether 'treatment' influences patient survival

c) Determine whether 'treatment' influences the development of graft dysfunction as assessed by presence of proteinuria (Protein:Creatinine Ratio > 50 or Albumin:Creatinine Ratio > 35) and change in estimated Glomerular Filtration Rate (eGFR).

d) Determine whether 'treatment' influences the rates of acute rejection in these groups

e) Determine the adverse effect profiles of 'treatment' in this group, in particular whether they are associated with increased risk of infection, malignancy or Diabetes Mellitus (DM).

f) Determine the cost effectiveness of routine screening for HLA Ab and prolonging transplant survival using this screening/treatment protocol.

g) Determine the impact of biomarker screening and "treatment" on the patients' adherence to drug therapy and their perceptions of risk to the health of the transplant.

Trial design including blinding

This is a prospective, open labelled, randomised marker-based strategy (hybrid) trial design, with two arms stratified by biomarker (HLA Ab) status. Recruitment will take place in 12 renal transplant units, recruiting for 45 months with recruits followed up intensively for at least 32 months (maximum 64 months) and primary endpoint assessed by remote evaluation after 43 months post-randomisation is achieved by all. The trial design is represented in the flow diagram in section 2.3, showing the number of patients anticipated to be in each group by the end of the trial, based on sample size calculations, consent rates, eligibility and estimated fall-out. Using the flow diagram (top-to-bottom) as a guide: recipients of cross-match negative transplants aged 18-75, > 1-year post-transplant with an eGFR ≥ 30 will consent to the screening/treatment process.

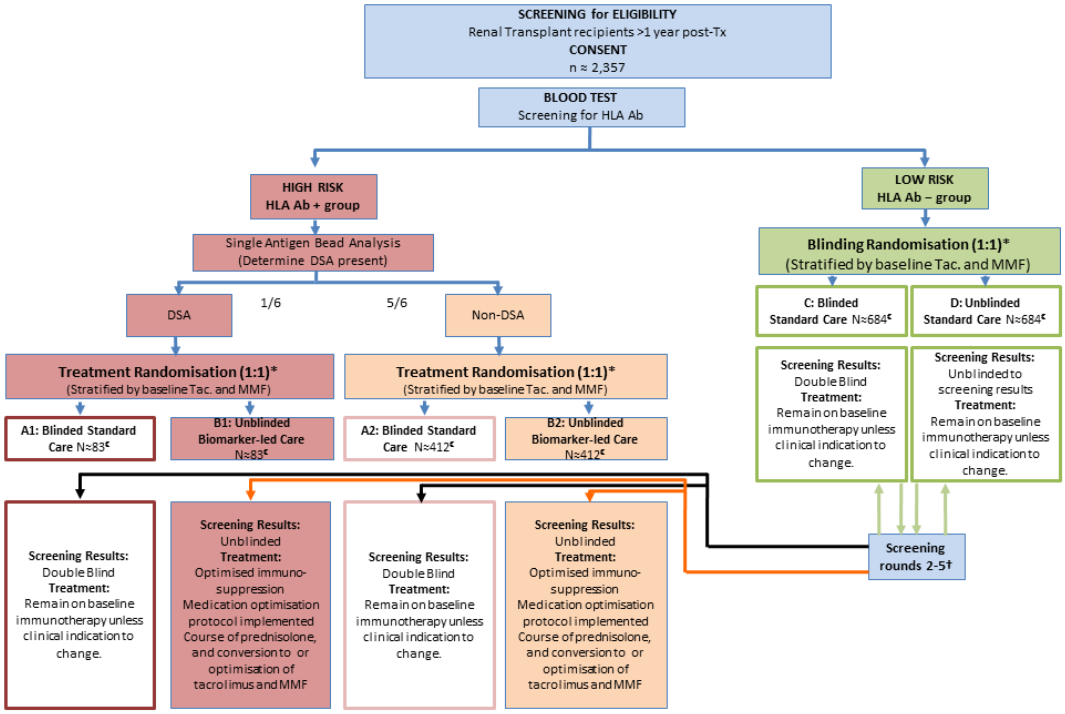
The first stratification will result from blood test screening for HLA Ab. Approximately 35% will be HLA positive, with ~65% negative. The HLA Ab+ patients will be further screened with single antigen beads to determine whether DSA are present (~1/6 DSA and 5/6 non-DSA). Thus, biomarker stratification leads to three groups (DSA+, non-DSA+ and HLA Ab-neg).

The second stratification will be based on current immunosuppression, to ensure balanced numbers already on Tac or MMF in each group. The final stratification will be by site.

HLA Ab positive patients will be randomized 1:1 into either Blinded Standard Care or Unblinded Biomarker led-care. Patients in the former (groups A1 & A2 in the flow chart in 2.3) will be blind to their biomarker status and will remain on baseline immunotherapy, whereas patients in the latter (groups B1 and B2 in the flow chart) will know their HLA Ab status and will be offered "treatment".

HLA Ab-negative patients will remain on their existing immunotherapy and randomized 1:1 into either Blinded (group C) or Unblinded groups (D), with only the latter knowing their HLA Ab status. Both these groups will receive regular Ab status monitoring for the first 3 years. Those patients who become positive during subsequent screening rounds (~10% per year) will be moved to the appropriate HLA Ab positive groups (DSA+ or non-DSA+) for final data analysis.

All patients in group D found to be positive on second or subsequent rounds will be offered the same "treatment" as those patients who were positive in the first screening round and be intensively followed up for an additional 32 months from the time they become positive. Thus, the maximum amount of time any single patient may remain in intensive follow up is 64 months[1]. New patients will be recruited to the study at each successive screening round.



*Randomisation performed on results of a recruit's first screening test. Those with HLA Ab undergo no further screening as part of the trial (but serum will be stored for analysis of HLA Ab profiles later). †Those initially HLA Ab-negative undergo routine screening every 8 months. THERE IS NO SECOND RANDOMIZATION: If a recruit allocated to Blinded standard care (group C) becomes HLA Ab positive (black lines), he/she remains in Standard care group (group A1 or A2). If in unblinded standard care group (D), they change to unblinded biomarker-led treatment care (group B1 or B2) (orange lines). € Numbers in each group are those anticipated at the end of study.

[1] For example, a patient recruited at the beginning of the study into groups C or D, found to have developed HLA Ab on the final screening round, will transfer into groups A or B and remain in intensive follow up for another 32 months.

Figure 1. Trial design flow diagram

Method of allocation of groups

Prior to randomisation but after consent, site staff will register all recruits online and each will be assigned a MACRO PIN. Samples from all recruits will be sent to the HLA laboratory, along with this PIN and a sample request form containing other information required for randomisation.

Laboratory staff will screen for HLA Ab and perform single antigen bead testing on positive screening samples to check for the presence of DSA. Once this information is known, the lab staff will access the randomisation system and randomise the patient, using the HLA Ab results and information on the sample form to stratify. Randomisation will be further stratified by centre and previous immunosuppression. In all sites, the PI's and nurses will be automatically emailed and the system will tell them whether the patient is in the blinded or unblinded groups. If in the unblinded group, it will feed back HLA Ab status to the PI. Unblinded patients will be identified by blue stickers to be appended to the notes and all future clinical samples. The system will tell the trial staff to enter HLA Ab-negative patients into the subsequent 8 monthly screening rounds, and also whether the patients have been selected to provide future samples for 4 monthly scientific analysis (for transfer to the CIs laboratory). This information will be relayed using a 'star' on the blue labels, appended to their laboratory request forms thereafter.

Blinded patients will have green stickers/labels. HLA Ab status will not be fed back to the PIs or trial staff. A 'star' will be used to tell the trial staff which recruits have been selected to provide 4 monthly samples for transfer back to the CIs lab for scientific analysis. All these patients will have samples taken 8 monthly for HLA Ab screening. Once inside the lab, the lab staff will use their knowledge of the HLA status to determine those from HLA Ab-negative patients which will undergo screening. The samples from HLA Ab positive patients will be discarded.

On the second and subsequent HLA Ab screening rounds, the lab staff will update the randomisation system. The results from patients in the unblinded groups only will be forwarded to the PI and laboratory staff, via email. This will indicate whether status has changed and trigger the initiation of the treatment protocol in those that have changed from HLA Ab negative to positive.

Randomisation will be via the online King's Clinical Trials Unit randomisation system.

Laboratory staff at each recruiting site, with access to HLA Ab results, will be provided with a unique username and password to access the randomisation system. Password access must be authorised by the trial manager in all cases and request directly from sites will not be processed. Access to the system is via https://cturandomisation.iop.kcl.ac.uk/OUTSMART, clicking the 'randomisation – advanced' link and selecting the OuTSMART Trial.

Duration of the treatment period

All treatments will be introduced on the basis that they will be tailored to the individual patient, according to compliance, tolerance and achievement of target levels (for Tac). Failure to tolerate one or more of the components of the protocol (or refusal to take any of the agents) will not be used as a reason for withdrawal from the study.

Frequency and duration of follow-up

Following recruitment to the trial, the patients will undergo 32 months of intensive follow up involving 8-monthly clinic visits post-randomisation, except in the following scenario; a patient in groups C or D who becomes Ab positive during the initial 32 months follow up will transfer to the relevant Ab+ group and undergo intensive follow up for a further 32 months from date of transfer. Therefore, the maximum amount of time that any single patient may remain in intensive follow up is 64 months.

The optimised treatment protocol will be introduced within the first 3 months in those HLA Ab+ patients allocated to this group. Recruits will be seen *up to* twice weekly during this period (maximum of 6 extra clinic appointments are envisaged), though they should be on maintenance dose prednisolone 7 weeks after initiating optimisation. During this period, they will have full blood count (as above), creatinine, Na^+, K^+, glucose, calcineurin inhibitor trough levels and blood pressure monitored according to the trial protocol. Once stabilized, they will be seen at least 8 monthlies in transplant clinic. Patients allocated to all other groups will be seen at least every 8 months in transplant clinic for formal study assessments. Patients may be seen at other times during this period, according to clinical need, but study assessments should be done within the time parameters established in the protocol.

Once every 8 months the following will be recorded. a) Weight and bp; b) Full blood count (minimum Hb, WCC, platelets); c) Biochemical series (creatinine, Na^+, K^+, bicarbonate, calcium, CRP, glucose); d) MDRD eGFR on latest creatinine; e) calcineurin inhibitor trough levels; f) protein: creatinine ratio on urine sample; g) total immunoglobulin levels; h) episodes of infection, malignancy or new DM; i) episodes considered to be adverse events. Every 16 months a lipid profile will be performed.

Once every 8 months, HLA Ab-negative patients will undergo further screening for HLA Ab (see above). At the end of the study, all patients will undergo a final test for HLA Ab status.

Upon completion of the 32- to 64-month intensive follow up, consisting of 8-monthly research visits, the participants will be told by the research nurses that they no longer need to attend to

clinic for research visits, but their regular clinic visits will continue. At the last intensive follow up visit a HLA Ab sample will be collected from all participants. The original plan was for the study to conclude once the final participant reached 43 months post-randomisation in June 2020 and in the three months prior to the conclusion of the study, data regarding the primary endpoint was to be collected from patient notes for all participants

Due to the coronavirus pandemic in the UK in 2020 most clinical trial activity, including this trial, was severely limited and so it was not possible to rely on the original plan to obtain the primary endpoint data between April and June 2020. For this reason, the best alternative was to obtain the primary endpoint data from patients' clinic notes once the coronavirus situation had eased and clinical trial activity re-started nationwide. Evidence for graft failure or death will now be taken from the participants' last hospital contact prior to March 16th, 2020. These data, which will reflect participants' pre-COVID status, will be used for the primary endpoint analysis. Evidence of graft failure or death will also be taken from participants' notes from their most recent hospital contact at the point of a final assessment between September 1st, 2020 and November 30th, 2020. During this designated three-month window, endpoint data will be collected from each patient's notes only once; the window is simply for pragmatic reasons to allow centre research staff sufficient time to check all patients' notes. These data, reflecting status post-onset of COVID crisis, will be used for a sensitivity analysis. The trial will conclude on November 30th, 2020.

Data collection

Eligibility Screening

Potentially eligible patients will be approached at a routine clinic appointment by the PI or research nurses and given printed and verbal information about the trial. They will have the opportunity to return for a second consultation within a few days to give informed consent for recruitment into the study or to do this on their next routine appointment. Alternatively, eligible patients will be sent information about the study through the post, for discussion and consent at their next routine appointment. Following consent, full eligibility criteria will be reviewed. This may include testing for chronic viral disease (if no such test within last 5 years) or pregnancy (if history suggests possibility of pregnancy).

Inclusion Criteria

- Sufficient grasp of English to enable written and witnessed informed consent to participate.
- Renal transplant recipients >1 year post-transplantation, male or female
- Aged 18-75 years
- Estimated glomerular filtration rate (eGFR by 4 variable MDRD) of ≥30 (within the previous 6 months of signing consent or taken at screening if not done in the previous 6 months).

Exclusion Criteria

• Recipient requiring HLA desensitisation to remove antibody for a positive XM transplant
• Recipient known already to have HLA antibody WHO HAS RECEIVED specific intervention for that antibody or for CAMR/chronic rejection
• Recipient of additional solid organ transplants (e.g. pancreas, heart, etc.).
• History of malignancy in previous 5 years (excluding non-melanomatous tumours limited to skin)
• HBsAg+, HepC IgG+ or HIV+ recipient (on test performed within previous 5 years)
• History of acute rejection requiring escalation of immunosuppression in the 6 months prior to screening.
• Patient enrolled in any other studies involving administration of another IMP at time of recruitment
The following exclusion criteria are based on information contained within the SMPcs of the IMPs
• Known hypersensitivity to any of the IMPs
• Known hereditary disorders of carbohydrate metabolism
• Pregnancy or breastfeeding females (based on verbal history of recipient)
• Pre-menopausal females who refuse to consent to using suitable methods of contraception throughout the trial.

Baseline

Post consent, patients who have not been screened for HIV or hepatitis B/C within the last 5 years will need to have additional screening tests for these viruses. Female patients who report they may be pregnant will have a blood test for beta-HCG levels. Once eligibility criteria have been met, the following baseline data will be recorded at recruitment:

a) Weight and bp;

b) Sex and ethnicity;

c) Age and date of birth;

d) HLA type and that of donor kidney (if known);

e) Any significant past medical history, including history of diabetes mellitus, cause of renal failure, details of previous transplants and cause of graft loss, evidence of sensitisation pre-transplantation (panel reactive antibody and antibody specificities if known);

f) medication list and doses;

g) PCR on urine sample

All patients will then have blood taken for;

i) Baseline clinical parameters: a) Full blood count (minimum Hb, WCC, platelets); b) Biochemical series (creatinine, Na^+, K^+, bicarbonate, calcium, CRP, lipid profile, glucose); c) MDRD eGFR on latest

creatinine; d) current calcineurin inhibitor 12 hour trough levels (as appropriate); e) Total Immunoglobulin levels

ii) Scientific analyses: 50-60 mls blood for separation of PBMC and 20mls for serum storage.

iii) Analysis of HLA Ab status (10mls clotted blood), as described above, which will allow randomization to proceed.

All patients will be asked to complete questionnaires to assess attitudes to risk/adherence.

Primary outcome measures

The primary endpoint is time to graft failure in HLA Ab positive patients randomized to biomarker led care groups vs. time to graft failure in HLA Ab + patients randomized to standard care groups assessed at a minimum of 43 months post-randomisation achieved by all (see above Section 1.5) for adjustment to final data collection following the coronavirus pandemic, consequently a few participants may not reach the minimum of 43 months for the primary endpoint. Graft failure will be defined as re-starting dialysis or requiring a new transplant.

Secondary outcome measures

The secondary clinical endpoints are:

• Time to graft failure in patients randomized to blinded HLA Ab screening vs those randomized to unblinded screening. Graft failure will be defined as re-starting dialysis or requiring a new transplant.

The following endpoints will be assessed at end of intensive follow up (32 months):

• Patient survival.

• Graft dysfunction, as assessed by two separate measures, presence of proteinuria (Protein Creatinine Ratio >50 or Albumin/Creatinine Ratio > 35) and change in estimated Glomerular Filtration Rates over 32 months.

• Rates of biopsy-proven rejection.

• Rates of culture- or polymerase chain reaction (PCR)-positive infection, biopsy-proven malignancy and DM.

• Health economic analysis of outcomes in intervention vs. control groups.

• Analysis of adherence and perceptions of risk in BLC groups.

Adverse Events

This trial fulfils the criteria for a 'Type A' trial (i.e. risk no higher than that of standard care) Therefore, there will be reduced reporting of adverse events. Definitions of expectedness reported below are based on those listed in the SmPC for each IMP.

Adverse Event (AE): Any untoward medical occurrence in a subject to whom a medicinal product has been administered including occurrences which are not necessarily caused by or related to that product.

Adverse Reaction (AR): Any untoward and unintended response in a subject to an investigational medicinal product which is related to any dose administered to that subject.

Unexpected Adverse Reaction (UAR): An adverse reaction the nature and severity of which is not consistent with the information about the medicinal product in question set out in the summary of product characteristics (SmPC) for that product.

Serious adverse Event (SAE): Serious Adverse Reaction (SAR) or Unexpected Serious Adverse Reaction (USAR): Any adverse event, adverse reaction or unexpected adverse reaction, respectively, that

Results in death;

Is life-threatening;

Required hospitalisation or prolongation of existing hospitalisation;

Results in persistent or significant disability or incapacity;

Consists of a congenital anomaly or birth defect.

Although not a serious adverse event, any unplanned pregnancy should be reported via the SAE reporting system as stated below.

Sample size estimation (including clinical significance)

The primary purpose of this trial is to demonstrate superior outcomes using a defined treatment strategy in biomarker (HLA Ab) positive patients and at the same time demonstrate non-inferior outcomes when the unblinded screening strategy is applied to the entire patient population. Time to graft failure has been chosen as a clinically relevant primary outcome. As a reference for power calculations, we have used the observed failure rates reported by Lachmann et al. [8] for HLA Ab+ and HLA Ab-neg patients. Since failure rates differ between DSA+ and non-DSA+ patients, sample size calculations have been carried out separately for these groups. Following these calculations,

we have estimated the number to be screened, based on expected drop out rates, expected screening results and eligibility criteria (see below).

We have based our estimates of the differences in primary outcome between groups on two things; first, the results of our preliminary data from patients with CR treated with a similar regime as used here; second, our assessment that large differences in primary outcome will be needed to make the screening programme cost-effective.

Hypotheses and power calculations:

1) Superiority on Biomarker Positive Patients:

1.1) A1>B1:

HLA Ab+ patients, with DSA, randomized to standard care (A1) will show higher graft failure rates than patients randomized to biomarker-led care (B1). We hypothesize that the experimental treatment will bring the failure rate in group B1 down to that of non-DSA patients in standard care (A2). Assuming that 30% of patients with DSA randomised to standard care (A1) will have experienced chronic rejection (CR) by 3-years follow up, we expect treatment optimisation to reduce the rate of CR in DSA patients randomised to group B1 down to 16% at 3-years follow up (rate observed in patients with non-DSA). This corresponds to a Hazard ratio (HR) of 0.489. The expectation is for 11% and 21% of CR among patients with DSA in in group A1 at 1 and 2-years follow up respectively (as in [8]), and extrapolating based on a HR of .489, we expect BLC to reduce those CR to 5.5%, and 10.89% at 1 and 2-years.

Using a variable follow up design assuming an average accrual monthly rate of 3.6 patients per month, and a minimum follow up time of 43 months, recruiting 165 patients with DSA would allow us to observe 23/83 (28%) events of CR in patients under biomarker led care (B1), and 39/82 (47%) in the standard care group (A1). This would provide 80% power and 5% type I error, for a two-sided log-rank test.

1.2) A2>B2:

HLA Ab+ patients, with non-DSA, randomized to standard care (A2) will show higher graft failure rate than patients randomized to biomarker-led care (B2). We hypothesize that the experimental treatment will bring the failure rate in group B2 down to that of biomarker negative patients in standard care (C). Assuming that 16% of patients with NDSA randomised to SoC will have experienced chronic rejection (CR) by 3-years follow up, we expect treatment optimisation to reduce the rate of CR in NDSA patients randomised to BLC down to 6% at 3-years follow up (rate observed in patients without HLA antibodies).

This corresponds to a Hazard ratio of 0.351. Based on Lachman et al. the expectation is for 3% and 11% of CR among patients with NDSA in SoC at 1 and 2-years follow up respectively, and extrapolating based on a HR of 0.351, we expect BLC to reduce those CR to 1.1%, and 4.1% at 1 and 2-years.

Using a variable follow-up design (patients followed until failure, drop out or end of minimum follow up), assuming an average accrual monthly rate of 15.5 patients per month, and a minimum follow up time of 22.4 months, recruiting 296 patients with NDSA, would allow us to observe 8/149 (5.3%) events of CR in patients under BLC, and 21/147 (14%) in the SoC group (total duration = 41.5 months). This would provide 80% power to determine a statistically significant difference between SoC and BLC, using a log-rank test, with a 2-sided type-I error rate of 5%.

The numbers enrolled in groups A & B include those patients initially enrolled in groups C or D who become HLA Ab+ during re-screening.

2) Non-inferiority of all Unblinded patients compared to all Blinded patients:

2.1) A1+A2+C ≥ B1+B2+D:

All patients randomized to unblinded screening will show equal or lower graft failure rates than all patients randomized to blinded screening, irrespective of biomarker status. At the end of the trial, we expect 58% of patients to be in the HLA Ab negative groups, 7% DSA+ groups and 35% non-DSA+ groups (after drop-outs). At the time of planning the OuTSMART study, we calculated that based on all assumptions above, all patients randomised to SoC combined would experience 13.9% of CR.

We established a non-inferiority limit of 5% absolute difference in rate of CR at 3-years, so that the BLC group would be considered inferior to SoC with a CR rate of 18.9% or higher (expectation under the null hypothesis). This corresponds to a HR of 1.4 under the null hypothesis, and a HR of 0.63 under the alternative. Recruiting 672 patients over a period of 13.2 months, at an average accrual rate of 51 patients per month, and a minimum follow up of 18.21 months, would allow us to observe 22/337 (6.5%) events of CR in the SoC group, and 32/335 (9.5%) in the BLC group.

This would provide 90% power to demonstrate non-inferiority with a one-sided 95% Confidence Interval of the HR estimated using a Cox regression model. Given the above proportions, this requires enrolling 336 patients in each of groups C and D and this should allow 423 total patients to reach the primary endpoint (i.e. remain negative (after dropouts) at the end of their three year follow-up).

Based on an overall expected proportion of 7% DSA participants (including from re-screening rounds) we will need to recruit 2357 patients overall to recruit the target of 165 DSA patients. Because of this requirement to recruit sufficient DSA participants, the recruits to the other groups are likely to be more than the minimum required for at least 80% statistical power for the individual hypotheses.

Brief description of proposed analyses

Analyses will be carried out by the trial statistician. In the first instance data will be analysed under intention-to-treat assumptions (i.e. analyse all those with data in groups as randomised irrespective of treatment received). As per Section 1.7, in addition using all participants, all outcomes will be analysed separately within the subgroups of HLA Positive DSA participants and HLA Positive Non-DSA participants.

Those patients who become positive during subsequent screening rounds will be moved to the appropriate HLA Ab positive groups (DSA+ or non-DSA+) for analysis. These participants will be included from the time they became HLA Ab positive (date sample for screening was taken). For the analysis using all participants, these participants will be included from the time of randomisation.

1 Further changes in procedures in response to Covid-19

Subsequent to the changes outlined in Section 0 (which detail changes made to primary endpoint collection following the COVID-19 pandemic), database lock was intended to be completed in February 2021, and the analysis completed by April 2021, with end date for OUTSMART the 30th April 2021. The remaining requirements prior to database lock taking place and analysis starting were i) that all data had been cleaned and ii) that final monitoring/source data verification of the primary endpoint had been completed as planned (100% of all graft failures, 25% of remaining participants). The monitoring is undertaken by King's Health Partners Clinical Trials Office (KHP-CTO). The KHP-CTO had intended to visit all sites to finish final monitoring of the primary endpoint (graft failure) data by the end of January 2021. However, monitoring was suspended as the latest wave of the COVID-19 pandemic occurred in Winter 2020/2021 and this meant that research nurse teams who were responsible for making source data available for monitoring in each centre were again redeployed. For these reasons, it was foreseen that the database lock could not be completed according to the planned timeline.

A 6 month no-cost extension to the project to 31st October 2021 was granted from the trial funder, National Institute for Health Research Efficacy and Mechanism Evaluation Programme (NIHR EME), with the intention that monitoring would re-commence once possible. However, the shift in timelines meant that the final cleaning and statistical analysis would be delayed which in turn impacted staffing, specifically the Trial Manager and Statistician, both essential at this key end stage of the trial but whose input could not be guaranteed through to the new study end. For these

reasons a new plan was devised that optimised the time that the trial manager and statistician had available:

- The database will be 'frozen' (temporarily locked) following completion of all data checking apart from the remaining KHP monitoring. This will allow the statisticians to run the analyses unblinded with the aim of completing them in April 2021 in keeping with the existing timeline.
- The dated statistical report will be generated from these analyses and sent to the DMEC so they can attest to the validity of the process and confirm they are happy for this report to be shared with the Chief Investigator.
- The Chief Investigator will then receive the report in order for the team to draft a preliminary Clinical Study Report and preliminary primary publication. These will not be finalised, however.
- Once KHP monitoring of the primary endpoint is able to be completed and any queries resolved, the database will be finally locked and the analyses re-run.
- It is expected that at most there would only be minor changes to the database following monitoring queries from the KHP-CTO, and so the results and their interpretation are very unlikely to change. It will be ensured that changes to the data are only made in response to monitoring queries raised by the KHP-CTO CRA. The MACRO EDC system has functionality to record any changes made between the database being frozen and final database lock and these will be checked by the KHP-CTO to confirm that all changes were made in response to monitoring queries only.
- The drafts of the preliminary clinical study report and preliminary primary publication will be updated with final results post database lock. The preliminary report results will be included as an appendix in both the final clinical study report and primary publication to ensure transparency.

Data analysis plan – Data description

Descriptives by HLA status

All descriptives will be broken down by HLA status. Where appropriate, descriptives for HLA Ab Positive participants may be further broken down by whether they were HLA Positive at randomisation or through re-screening. Those who became HLA Ab Positive at re-screening may also be described separately prior to becoming HLA Ab Positive and after becoming HLA Ab Positive.

Recruitment and representativeness of recruited patients

CONSORT flow chart will be constructed (1) – see Figure 2. This will include the number of eligible patients, number of patients agreeing to enter the trial, number of patients refusing; number of HLA Ab positive and negative patients, then by treatment arm: the number continuing through the trial, the number withdrawn or lost to follow-up and the numbers excluded/analysed.

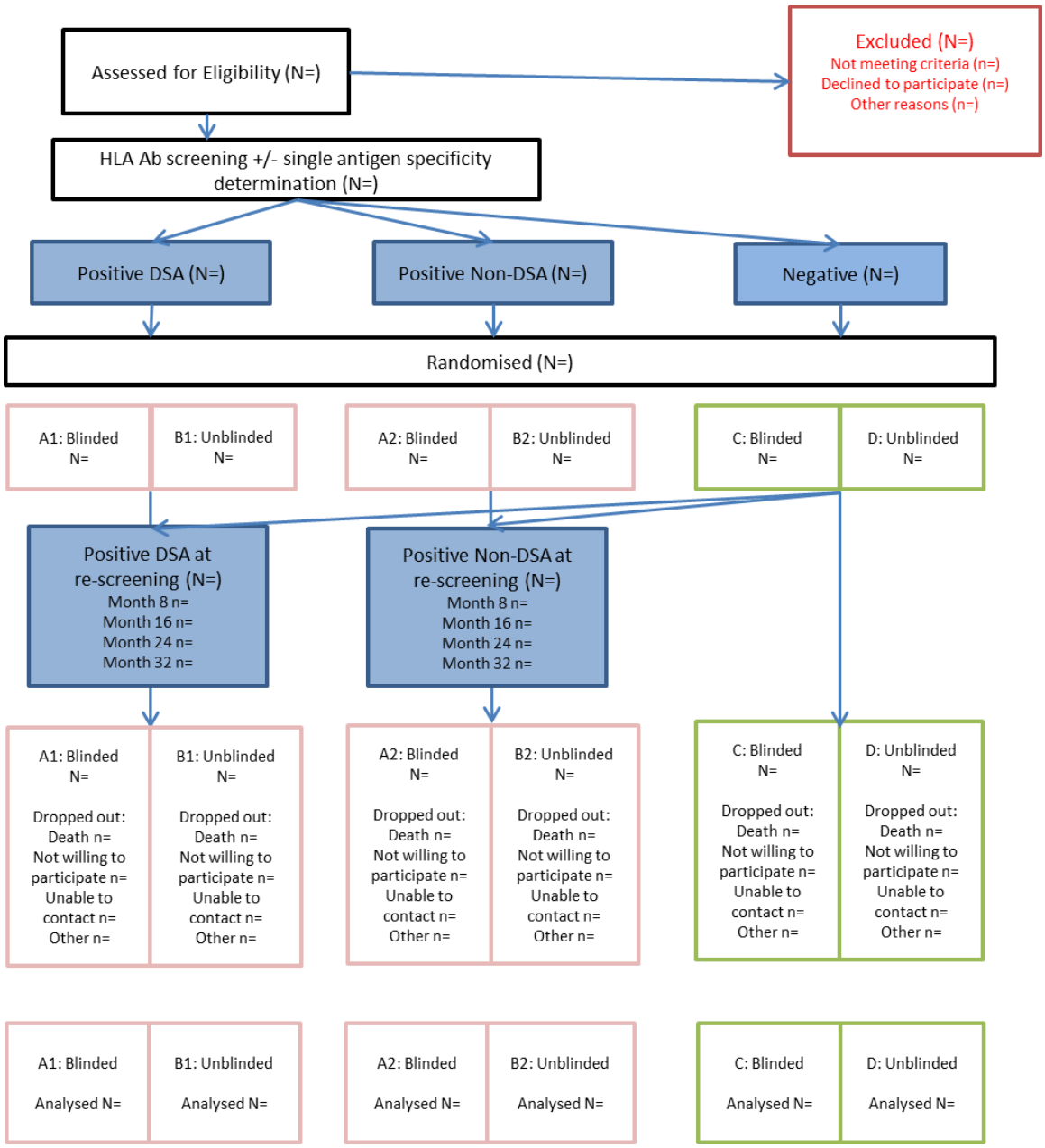


Figure 2. Template CONSORT diagram for OuTSMART trial

Baseline comparability of randomised groups

Baseline descriptions of participants by trial arm within HLA status and overall: means and standard deviation or numbers and proportions as appropriate. Especially relevant factors are age, sex, HLA mismatches, baseline eGFR, site, previous treatment, and time from transplantation. No significance testing will be performed.

Adherence to allocated treatment and treatment fidelity

Antibody positive patients randomized to biomarker led care who take the full optimization protocol (three drugs) will be compared to those who don't take the full protocol (less than three drugs) with respect to baseline variables and outcomes.

Loss to follow-up and other missing data

The proportions of participants with missing data for each variable will be summarised in each arm and at each time point. The baseline characteristics of those missing follow up will be compared to those with complete follow up. The reasons for withdrawal from the trial will be summarised.

Adverse event reporting

Adverse events (AE), adverse reactions (AR), serious adverse events (SAE) and serious adverse reactions (SAR) will be summarised by group (within trial arm and HLA status) as proportions and 95% confidence intervals.

Assessment of outcome measures (unblinding)

Evidence for unblinding of HLA status (within blinded arm) will not be assessed.

Descriptive statistics for outcome measures

Frequencies and descriptive statistics of all primary and secondary outcome measures will be reported with estimates and 95% confidence intervals.

Description of therapists/therapies

Immunosuppressive drugs and doses prescribed to each group of patients will be summarized.

Data analysis plan – Inferential analysis

Main analysis of treatment differences

Analysis of primary outcomes

Statistical analysis will be on an intention to treat basis. Patients who become positive during the follow up will be included in the appropriate group (group labels refer to flow diagram in figure 1).

1. Superiority:

H_0: $h_{A1}(t) = h_{B1}(t)$ & $h_{A2}(t) = h_{B2}(t)$

H_1: $h_{A1}(t) \neq h_{B1}(t)$ & $h_{A2}(t) \neq h_{B2}(t)$

In order to test superiority for the primary outcome in the Biomarker (HLA Ab) positive groups (Hypothesis 1.1 and 1.2), we will use Cox proportional hazards regression models to estimate the graft failure hazard ratio between the biomarker led care and standard care groups and test at the 5% level of significance. Results will be given as estimates and 95% CIs. Within the model, we will adjust for previous immunosuppression regimen and research site (as these are the randomisation stratification factors) for increased statistical efficiency.

We will check the proportional hazards assumption by examining Kaplan-Meier plots and by testing for an interaction between group (BLC or SC) and time to graft failure within the model.

2. Non-inferiority:

H_0: $h_{Unblind}(t) / h_{Blind}(t) \geq \delta$

H_1: $h_{Unblind}(t) / h_{Blind}(t) < \delta$

In order to test for non-inferiority of the unblinded groups compared to the blinded groups (hypothesis 2.1), we will use Cox proportional hazards regression models to estimate the graft failure hazard ratio. We will adjust for the stratification factors in the model as outlined above and check the proportional hazards assumption by examining Kaplan Meier plots and by testing for an interaction between unblinded/blinded group and time to graft failure. We will conclude non-inferiority if H_0 gets rejected at 5% significance, and the corresponding upper bound of the 95% CI for the hazard ratio excludes the limit δ (hazard ratio of 1.4).

Analysis of secondary outcomes

All secondary outcomes will be analysed comparing standard care versus biomarker led care groups within the HLA positive DSA participants and within HLA positive non-DSA participants as well as between unblinded and blinded groups overall, as per the primary outcome analysis. We will use a similar procedure using Cox proportional hazards regression for the analysis of secondary time to event (survival) outcomes. Where numbers allow, secondary binary outcomes will be analysed using logistic regression with adjustment for stratification factors. Where numbers are too small for this, the Z-test or Fisher's exact will be used. For continuous secondary outcomes, linear regression will be used (or linear mixed models where accounting for repeated measures), adjusting for baseline values of the outcome and stratification factors. Transformations will be considered where data is skewed.

Results will be given as estimates (odds ratios or differences in proportions) and 95% CIs.

Statistical considerations

Missing baseline data

Missing baseline data should not be an issue for the primary analysis. Some extensions to this analysis may use other baseline variables; if these contain missing data, the number with complete data will be reported and they will be imputed using a method suitable to the variable as per the recommendations of White and Thompson (2).

Missing outcome data

If post treatment variables such as compliance with treatment are found to be predictive of drop out, multiple imputation will be considered.

Method for handling multiple comparisons

No formal adjustment of p-values for multiple testing is necessary. However, care will be given to the interpretation of inference for the numerous secondary outcomes.

Method for handling non-compliance (per protocol/CACE analyses)

An exploratory per protocol analysis will be carried out comparing time to graft failure in only those participants who were optimised to the full treatment protocol (as defined in Section 3.1 of the Trial Protocol) in the unblinded arm against all blinded participants, within both the HLA Ab+ DSA and HLA Ab+ non-DSA groups. These analyses will be clearly stated as exploratory in the primary paper/report and will be interpreted accordingly.

If there are concerns about compliance, complier average causal effect (CACE) analysis will also be performed.

Model assumption checks

The proportional hazards assumption will be checked for the primary outcome model by testing for an interaction with time. For secondary outcomes, where we assume normally distributed outcomes; this will have been checked when describing the data and if substantial departures from normality occur, transformations will be considered. Residuals will be plotted to check for normality and inspected for outliers.

Sensitivity analyses

Due to the COVID-19 pandemic, the original primary outcome endpoint was moved forward and will be collected from notes to prior to March 2020 (see section 1.5). This unavoidable protocol change may lead to a slight reduction in statistical power for the primary analysis. A sensitivity analysis will be carried out for the primary outcome using, additionally, data from participants' most recent hospital contact as of the assessment period between September 1st, 2020 and November 30th, 2020. This sensitivity analysis may have increased power to detect a difference due to the longer observation time but might be biased due to the effects of the pandemic on the study participants. The sensitivity analysis will otherwise be carried out in exactly the same way for each of the hypotheses (for the primary outcome only).

Planned subgroup analyses

An exploratory subgroup analysis will be carried out within the HLA Ab+ non-DSA group using only those participants within this group that have definite non-DSA (as opposed to those participants who are classified as non-DSA as it is unknown whether they had DSA) This analysis will be clearly stated as exploratory in the primary paper/report and will be interpreted accordingly.

There are no other planned subgroup analyses beyond those described by the stratifiers.

Exploratory analyses

There will be no other exploratory analyses in the primary paper.

Exploratory mediator and moderator analysis

There will be no exploratory mediator or moderator analyses in the primary paper and so these will not be covered here; subsequent papers may explore this.

Interim analysis

There are no planned interim analyses of the primary outcome, but the statistician will carry out descriptive analysis of demographic variables, adverse events and other relevant factors to report during DMC meetings. The proportions of patients presenting HLA antibodies will be closely monitored to check that assumptions on which sample size calculations are made hold.

Software

Data management: An online data collection system for clinical trials (MACRO; Elsevier) will be used. This is hosted on a dedicated server at KCL and managed by the KCTU. The KCTU Data Manager will extract data periodically as needed and provide these in comma separated (.csv) format.

Statistical analysis: Stata version 15.1 and/or R will be used for all statistical analyses.

B) ECONOMIC ANALYSIS AND ADHERENCE/RISK ASSESSMENT PLANS

Heath economic objectives

Economic measures

Effectiveness: The economic evaluation will adopt an NHS perspective. Costs of intervention include the cost of screening beads and enhanced drug costs. Other costs will be calculated by combining service use data measured with the Client Service Receipt Inventory and hospital records where available with appropriate unit cost information. Costs will be compared between groups using bootstrap methods to take account of likely skewed data distributions. Costs will be combined with clinical outcomes ((1) graft failure; (2) patient survival (3) graft dysfunction (see defn. above); (4) acute rejection; (5) culture-positive infection, malignancy or diabetes) and (6) QALYs derived from the EQ-5D, a health-related quality of life measure Incremental cost-

effectiveness ratios will be computed to indicate the extra cost incurred to achieve an extra unit of outcome. Uncertainty around these ratios will be assessed using cost-effectiveness planes. Net benefit per patient will be calculated by multiplying QALY gains by the assumed maximum willingness-to-pay for a QALY (£20,000) and subtracting costs. Alternative values for a QALY will also be used and a cost-effectiveness acceptability curve derived.

Adherence and Risk Assessment.

Perceived risk of transplant failure will be measured using a specifically adapted version of the illness perceptions questionnaire (IPQ-brief). Inferential statistics (independent t-test if data is approximately normally distributed, Mann-Whitney U test otherwise) will be used to determine if a difference exists between Ab positive unblinded and blinded participants. Differences between the blinded and unblinded HLA Ab screening groups will also be investigated.

Participants' report of adherence behaviour will be assessed using the Medicines Adherence Report Scale (MARS), a valid and reliable scale that has been previously used to assess adherence in renal transplant recipients. Self-report measures have the advantage of being inexpensive and non-intrusive. However, it is known that self-report underestimates the true extent of nonadherence because of inherent self-presentational and recall biases. Self-presentational bias occurs when respondents may be reluctant to admit to nonadherence because they perceive a social contract where the expectation is one of high adherence. The MARS takes steps to diminish this bias by sanctioning and normalising reports of nonadherence. However, this does not totally remove the effect self-presentational and recall biases that are inherent in all self-report measures.

For this reason, we will apply a combined approach to adherence assessment, where initial categorisation of patients into high vs. low on the basis of self-report is revised based on calcineurin inhibitor (CNI) blood monitoring (carried out in routine management for patients prescribed tacrolimus or ciclosporin). In this approach, reports of low adherence are accepted as self-presentational biases act in the opposite direction (reports of low adherence are more reliable than reports of high adherence). Patients who report high adherence are reclassified to low adherence on the basis of CNI results (e.g. if levels are undetectable then participant is assumed to be nonadherent)). Inferential statistics (independent t-test if data is approximately normally distributed, Mann-Whitney U test otherwise) will be used to determine if a difference exists between the adherence to immunosuppressive medication by Ab positive unblinded and blinded participants.

In order to explore the potential antecedents to participants' adherence behaviours, they will also be asked to complete specially adapted versions of questionnaires relating to treatment intrusiveness (TIQ), symptoms associated with immunosuppressants (SAQ), beliefs about

medicines (BMQ), and whether they are feeling anxious and/or depressed (HADS). Inferential statistics (independent t-test if data is approximately normally distributed, Mann-Whitney U test otherwise) will be used to determine whether systematic differences exist between Ab positive unblinded and blinded participants. Differences between the blinded and unblinded HLA Ab screening groups will also be investigated.

Adherence and perceptual data will also be linked with clinical outcome data (graft failure, patient survival, graft dysfunction, acute rejection, culture positive infection, malignancy or diabetes) in order to determine any relationships. A regression model will be used for this analysis.

All patients taking part in the trial will be asked to complete the questionnaires at baseline and then at 24 months. Participants in whom the 'clock is reset' will also complete the questionnaire 24 months after this point in time. Questionnaires will be administered electronically.

C) SCHEDULE OF ASSESSMENTS AND MEASURES

Phase	Peri-Randomization	Post-Randomization										
		Unblinded HLA Ab+ groups – Approximate times of assessment (+/- 1 week). Once stabilised, go to month 4 assessment						All Groups – Approximate times of assessment (+/- 3 months)				
Study Week/month	Day -56 to 0	Wk 2	Wk 4	Wk 6	Wk 8	Wk 10	Wk 12	Month 8	Month 16	Month 24	Month 32	Min 43 months
Informed consent	x											
Inclusion/Exclusion Criteria	x[2]											
Medical History inc. Drugs	x[3]											
Transplant / sensitisation Hx	x											
Registration / Demographics	x[4]											
Weight / BP	x							x	x	x	x	
Urine PCR	x							x	x	x	x	
Haematology[5]	x		x		x		x	x	x	x	x	
Biochemistry	x[6]		x[7]		x[8]		x[8]	x[8]	x[9]	x[9]	x[10]	

[2] Including virology and pregnancy testing where appropriate.
[3] For registration, need to know whether already on tacrolimus and / or MMF/myfortic.
[4] Do this prior to taking blood for HLA Ab screening
[5] Hb, WCC, platelet count at all time periods
[6] Creatinine, Na^+, K^+, bicarbonate, calcium, CRP, lipid profile, glucose, HbA1c.
[7] Creatinine, Na^+, K^+, glucose, HbA1c
[8] Creatinine, Na^+, K^+, bicarbonate, calcium, CRP, glucose, HbA1c
[9] As enrolment biochemistry

[Calcineurin inhibitor] trough	x	x[10]	x	x	x	x	x	x	x	x	x	
Total immunoglobulin (or IgG, IgM +/- IgA)	x								x		x	
HLA antibody screening	x[11]							x[12]	x[12]	x[12]	x[12]	
Samples for scientific analysis	x[12]							x[13]	x[13]	x[13]	x[13]	
Apply optimized treatment protocol[13]		x	x	x	x	X	x					
See Trial-specific Nurse	x							x	x	x	x	
Record Medications	x							x	x	x	x	
Adverse Events Form		x	x	x	x	X	x	x	x	x	x	
Questionnaire for analysis of adherence / risk	x									x		
Questionnaire for health economics	x								x			
Primary Endpoint (remote data collection)												

[10] In those patients having optimization of tacrolimus – continue until trough levels achieved
[11] At enrollment, on everyone. Beyond enrollment, send sample from recruits in unblinded HLA Ab-negative group and ALL blinded patients.
[12] At enrollment, on everyone. Beyond enrollment, only on those identified by a * 'star' on trial documents / labels. Collection of these samples will continue throughout the trial as long as there are resources available.
[13] Ideally participant will see a physician once a month whilst being optimized. Visit details are recorded in an Optimisation Log and not in the eCRF.

D) Reference list

(1) Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001 Apr 14;357(9263):1191-4.

(2) White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med 2005 Apr 15;24(7):993-1007.

E) Amendments to version 1.0

Version 2.0:

Primary endpoint was changed from graft failure rates at 3 years to time to graft failure as per OUTSMART protocol V11 26_11_2015. Changes were made to this document throughout to reflect this; the sample size calculation was amended, and the primary analysis was changed to use a Cox proportional hazards model. Changes were also made to reflect the change as per the protocol of 8-monthly follow up visits instead of 4-monthly follow up visits.

Version 2.1:

Changes were made to the definition and analysis of the graft dysfunction secondary outcome. The exploratory analyses were clarified. Amendments were made to the Adherence and Risk assessment analyses. Minor changes to wording to bring in line with OUTSMART protocol.

Version 2.2:

Clarified in some sections the three separate groups being analysed (HLA +ve DSA, HLA +ve NDSA and overall). Added more information on descriptives by HLA Status for participants who became HLA +ve through re-screening. Other minor changes to wording as per DMC comments.

Version 2.3:

Amendments made to primary analysis following the COVID-19 pandemic; change to final data collection of the primary endpoint and addition of sensitivity analysis for later data collection.

Version 2.4:

Amendment to clarify change to analysis timing (pre and post data lock) following impact of COVID-19 pandemic on KHP-CTO monitoring of primary endpoint and subsequently database lock (see section 1.8.1).