Development of an objective, standardised tool for surgical assessment of deceased donor kidneys: The Cambridge Kidney Assessment Tool

Authors: John OO Ayorinde[1] MBBS, Mazin Hamed[1] PhD, Mingzheng Aaron Goh[1] MD, Dominic M Summers[1] PhD, Anna Dare[1] PhD, Yining Chen[2] PhD and Kourosh Saeb-Parsy[1] PhD.

Affiliations: 1. Department of Surgery, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 0QQ, United Kingdom.

2. Department of Statistics, London School of Economics, Columbia House, Houghton Street, London, WC2A 2AE, United Kingdom.

Running title: Standardising surgical assessment of deceased donor transplant kidneys.

Corresponding author: Mr Kourosh Saeb-Parsy, Level 9 Department of Surgery, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 0QQ, United Kingdom. Email: ks10014@cam.ac.uk | Tel: 07713641416

Abstract: Quality assessment in kidney transplantation involves inspection to identify negative markers of organ-quality. However, there is a paucity of evidence guiding surgical appraisal, and currently there is no evidence to differentiate important features from those that can be safely ignored. We propose a method to standardise surgical assessment and derived a simple rule to rapidly identify kidneys suitable for transplantation.

Donor and recipient data were recorded alongside clinical outcomes in a prospectively maintained database. We developed a proforma (Cambridge Kidney Assessment Tool, CKAT) and used it to assess deceased-donor kidney transplants. Factors predictive of utilisation were identified by multivariate and univariate logistic regression analysis of CKAT-assessment scores, and test performance was evaluated using standard 2x2 contingency tables.

97 kidneys were included at a single centre (2013-2014), 184 CKAT-assessments were performed. A CKAT-threshold of 'Carrell+Perfusion>3', was highly specific (99%) and performed favourably to consultant opinion (specificity 95%). 96% of the kidneys implanted in accordance with the rule survived to 1-year (mean eGFR 45.3ml/min/1.73m^2).

To our knowledge, this is the first attempt to objectively define macroscopic features that are relevant to kidney utilisation. Common language could support training in organ assessment and ultimately help address unnecessary discard of donor kidneys.

Key words: Kidney, Macroscopic assessment, Surgical assessment, Deceased donor transplant, proforma, utilisation.

ABBREVIATIONS

CKAT	Cambridge Kidney Assessment Tool
DBD	Donation after brain death
DCD	Donation after circulatory death
ESKD	End Stage Kidney Disease
NHS	National Health Service
NHSBT	NHS Blood and Transplant
NPV	Negative Predictive Value
PPV	Positive Predictive Value

INTRODUCTION

Kidney transplantation is the gold standard treatment for End Stage Renal Disease, yet in the UK, patients wait an average of 3 years for a transplant. During this time, 12% of registered patients either become too unwell or die before they have their operation[1]. Despite this, 10-12% of donor kidneys are deemed unsuitable for transplantation following assessment by clinicians, and estimates for organ discard are even higher in the USA[2,3]. Whilst a small proportion of discards are due to an absolute contraindication to transplantation (e.g. malignancy), four of the top five reasons for discard are due to the team's assessment of organ quality[2]. Currently, quality assessment focusses on visual inspection to identify macroscopic features (e.g. perfusion characteristics) that, if present, are presumed to denote poorer quality grafts. Whether these features in fact have deleterious effects on transplant function has not been demonstrated conclusively. To complicate matters, organ assessment is not standardised, decision making - whilst challenging - is opaque, and therefore difficult to evaluate by external observers. With the introduction of the 'Fast-track' system in the UK, we now know that kidneys declined by five or more independent assessors can still be implanted, with comparable short-term results to standardly allocated organs[4]. Similarly, for unilateral kidney donors in the USA (where the paired kidney was declined by all transplant centres), kidneys still can provide good transplant outcomes[5]. Kidney assessment technologies (e.g. pre-implantation biopsy or ex-vivo machine perfusion) could enable clinicians to make more reliable assessments of organs after they are retrieved, but definitive evidence that these tests improve either the quality or numbers of kidney transplants is not yet at hand[6,7]. In any case, such technologies are not always widely available, and even when accessible, their use incurs additional time, logistical and financial costs.

Over the last 20 years, clinicians have gained access to large multicentre electronic databases, revealing population-based associations between donor risk factors (e.g. age, cardiovascular morbidity) and poorer transplant outcomes[8,9]. In comparison, there is a paucity of evidence guiding the surgical appraisal of kidneys for transplantation, and currently there is no evidence base to differentiate factors that should be relevant to utilisation decisions from those that can be safely ignored. The potential tools for appraisal of deceased donor kidneys were recently reviewed, and macroscopic assessment was identified as an area urgently requiring research to address the variability and validity of organ assessments[10]. Disturbingly, whilst the surgeon's appraisal has a direct influence on organ utilisation, we have no evidence that independent assessors concur on assessment of organ quality parameters. A further complication is the fact that macroscopic appearances can evolve during the retrieval process, or during transport. Consequently, recipient centres often report discrepancies between their assessments of the organ after it arrives at their centre, with that of the retrieving surgeon(s). Significant discrepancy can mean that an organ is deemed unsuitable for its intended recipient and discarded. Although the extent of this problem is not well defined, as an indication, Callaghan and colleagues re-evaluated 20 discarded kidneys in the UK to determine if local assessing surgeons agreed with the decision to discard a kidney made by another centre. The assessors disagreed with the decision to discard in 65% of cases, and the most common reason for discard in that series (15 of 20 kidneys) was "poor perfusion"[2]. This suggests that inter-observer variability in assessment may contribute to misallocation, and/or unnecessary discard of kidney transplants.

A possible solution is for surgeons to utilise agreed parameters for description of deceased donor kidneys. Aside from reducing inter-observer variability, a common

language would support training in organ assessment, aid reporting and comparative studies, and could be used as the basis for an objective appraisal tool. The ideal assessment tool would be quick and easy to use, have good intra- and inter-observer concordance and should not rely significantly on user experience. As a tool to support organ selection, it should also have high positive and negative predictive values. If effective, organ utilisation could be rationalised without significant financial, logistical or time costs, and the tool could complement other assessment and allocation methods. This would be particularly beneficial for units that, for example, find the logistics of ex-vivo organ perfusion a challenge, or where there is limited access to a dedicated, integrated on-call histopathology service.

Therefore, this study had two objectives. First, we aimed to move towards standardised language in the appraisal kidneys for transplantation by producing an objective, standardised assessment tool for surgeons to use whilst assessing organs for transplantation. We have named this the Cambridge Kidney Assessment Tool or CKAT. Second, we aimed to determine which macroscopic features are most relevant to utilisation decisions and use these to provide a simple assessment method that can support decision making in transplant kidney utilisation.

Ethics

Approved by Cambridge University Hospitals NHS Foundation Trust, Service Evaluation Number 2324. This study complies with the Declaration of Helsinki (2013).

MATERIALS AND METHODS

Retrieval and allocation

During the study period (2013-2014), 97 deceased donor kidneys were retrieved, accepted, and assessed for transplantation at our centre. Kidney procurement was performed as described previously[11]. At the time of entry, kidneys from donors who had suffered irreversible, catastrophic brain injury sufficient to meet brain-death criteria (DBD), were allocated nationally by NHS Blood and Transplant (NHSBT), using an algorithm that incorporated HLA matching, time on the waiting list, level of sensitization to HLA and donor-recipient age difference[12]. Kidneys donated after circulatory death (DCD) were from Maastricht Category III donors[13], and donation was pursued for 4 hours following the withdrawal of life sustaining treatment. At the time of the study, a national sharing scheme had not yet not been implemented for DCD kidneys, and at least one kidney was retained by the local centre for a local recipient. Post-transplant care and immunosuppression were administered according to standard protocols, described more fully elsewhere[14].

Assessment process

During the study period, absolute contraindications for acceptance of a kidney transplant were: active IV drug use at the time of offer and active HIV infection. Relative contraindications included: malignancy, age and cardiovascular co-morbidity. Once the kidney arrived at our centre, kidneys from donors aged 60 years and over were biopsied and a Remuzzi score was assigned following histopathological assessment[15]. At the time of the study, kidneys scoring 1-3 were used as single transplants, 4-6 as duals, and those scoring 7 or more were discarded[16–18]. During preparation on the backbench, the implanting surgeon would

inspect the appearance of the kidney and evaluate its suitability for transplantation by integrating the results of the macroscopic assessment with all the information available from the donor's clinical history and events that occurred during retrieval process.

Appraisal method

We consulted a group of experienced transplant surgeons at our centre to identify the range of macroscopic features relevant to organ utilisation decisions. An assessment proforma was devised which included all the relevant features [fig 1], and the proformas were made available to surgeons performing organ retrieval and backbench preparation of kidneys destined for implantation at our centre. Prior to the determination of suitability for transplant made by the implanting surgeon, retrieval surgeons were asked to assess the kidney individually, using the assessment proforma[fig 1]. They were therefore unblinded to the clinical history, donor characteristics, and operative findings; mirroring clinical practice[fig 2]. A small number of kidneys were subjected to pre-implantation biopsy assessment, however, the results of these tests were not known at the time of the proforma assessment. The suitability for transplantation was also assessed by an independent transplant consultant at our centre who was not part of the implanting team and who did not use the proforma. This consultant was also unblinded to the clinical history, donor characteristics, and operative findings, mirroring clinical practice. A summary of the study appraisal pathway is given in Figure 2[fig 2].

Statistical analysis

Donor and recipient demographic data were recorded alongside clinical outcome data in a prospectively maintained database. Descriptive data were presented as number, median or percentage (categorical variables). Comparisons were made using the Kruskal-Wallis test (continuous variables), at a significance threshold of less than 0.05 for the p-value. Factors important for predicting utilisation of transplant kidneys were identified using univariate and multivariate logistic regression analysis of the assessment scores, described more fully below. Evaluation of the appraisal methods (i.e. Consultant vs CKAT-based appraisal) employed standard statistical approaches to compare test performance: sensitivity; specificity; positive- and negative-predictive values. Estimated glomerular filtration rate (eGFR, mL/min/1.73m2) was calculated using the 4variable MDRD formula at 1-year[19]. Graft survival is defined as time from transplant to graft failure (need for dialysis) censored for death with a functioning graft. Statistical tests were implemented using R Studio version 3.5.1 (2018) with the 'tidyverse' and 'caret' libraries installed as well as associated dependent packages.

RESULTS

Donor and assessor demographics

97 kidneys were included in this study, of which 90 were donated after circulatory death (DCD) whilst the remainder (n=7) were donated following the diagnosis of brain stem death (DBD). This reflects the fact that at the time of the study, at least one kidney from a DCD pair was retained by the local kidney transplant centre. Kidneys that were declined as unsuitable for transplantation prior to the arrival of the kidney back at base could not be included, as we required 3 independent assessments (retrieval surgeon, implanting surgeon and independent consultant) in our analysis. Exclusion from further analysis was therefore on the basis of: donor-anatomical factors (n=1, severe pre-mortem aortic dissection), damage (n=0) malignancy or infection (n=2).

In the UK the Consultant is a senior surgeon who has overall responsibility for the care of patients in hospital. The equivalent surgeon grade in the United States is Attending Surgeon. In the UK, a trainee is doctor who is pursuing a career in surgery. This includes those who are engaged in an official training programme, either at the core- (CT1-3), or specialty-trainee level (ST3-8). The workforce in Cambridge is international and welcomes many graduates from outside the UK whose training pathways do not map directly onto ours, as well as UK graduates outside the official training pathway. In this study, we have divided the grade of 'trainee' into junior trainee (3-6 years post-graduation) and senior trainee (>6 years post-graduation) which better characterises the important distinctions in our unit. The equivalent training grade in the United States is Residency. 93% of initial proforma assessments in this study were performed by junior and senior trainees,

reflecting the staffing of the Cambridge organ retrieval team (NORS team) [tab 1B]. Summary statistics regarding donor demographics can be found in Table 1A [tab 1A], which primarily reflects the DCD donor pool at the time of the study.

Assessment scores

184 assessments were made in total [tab 1B]. The median number of assessments per kidney was 2, and the maximum number of assessments was 4[fig 3]. The distribution of the scores for each assessment dimension can be seen in Figure 4[fig 4]. The actual number of assessments for each kidney depended on the availability of the staff at the time of retrieval. If scores between assessors differed, we used a majority vote to determine the "aggregated" score in our analysis and in the event of a draw, we took the higher score as a cautious measurement. Other ways of aggregating scores are also possible, but sensitivity analysis did not reveal significant changes in results when these were applied.

Inter-observer scoring using the proforma was highly consistent; in each dimension, no two assessors differed in their scoring of the kidney by more than 1 point on the rating scale.

Statistical methods used to develop the assessment tool

A two-stage decision procedure was proposed. In the first stage, we discarded all kidneys which were unusable prior to assessment due to donor-anatomy, damage, malignancy or infection (n=3) [fig 5]. In the second stage, we examined the recorded assessment scores. First, a univariate logistic regression analysis was used to identify the assessment characteristics that predicted utilisation, which then formed the basis of an assessment rule to guide organ selection. Of all the factors included in the proforma, three contributed to the predictive power of the logistic regression

model. They were quality of Carrell patch, extent of renal artery atherosclerosis and the kidney's perfusion characteristics. We then applied multivariate logistic regression based on these scores. Carrell patch and perfusion characteristics shared similar coefficients, while renal artery atherosclerosis no longer conferred additional predictive power in the presence of the aforementioned factors. Consequently, we removed renal artery atherosclerosis and combined the Carrell patch and the perfusion characteristics into one new factor.

This led to the decline rule of "Carrel patch quality + Perfusion quality > a threshold".

Determining the threshold and testing the rule

To determine the appropriate threshold for decline of a kidney for transplantation, various rule thresholds were retrospectively applied and their performance compared against the implanting surgeon's decision as the 'gold-standard'. A 2x2 contingency table was constructed to demonstrate sensitivity, specificity, positive and negative predictive values for each threshold. Based on this analysis, the best performing rule threshold for decline of a kidney for transplantation was: "Carrel patch quality + Perfusion quality >3". A higher or a lower threshold would worsen the sensitivity or the specification of the rule. To allow comparison of performance with current standard procedure, we also compared the accuracy of chosen threshold rule (Carrel + Perfusion >3) to the judgement of an independent, unblinded transplant consultant from our centre who was asked to assess the kidney and decide whether they thought the kidney was transplantable.

Assessment tool performance

Of the 94 kidneys that were included in the analyses, 73 were transplanted. Of these, 6 were implanted as part of a dual transplant, with the remainder implanted

singly [fig 5]. In cases where the Aortic patch was not used but the kidney was still transplanted, the anastomosis was end-to-side with interrupted sutures. Of the 21 kidneys declined for transplantation by the implanting surgeon, in 7 cases the prediction of the independent consultant (sensitivity 33%) matched this outcome, whilst 9 were identified by the assessment tool (sensitivity 47%), these kidneys were not accepted for transplant by any other centre and were discarded. Of the 73 kidneys that were ultimately transplanted, 69 were predicted by the independent consultant (specificity 95%), whilst 72 were identified by the assessment tool (specificity 99%).

Of the 11 kidneys declined for transplantation by the independent consultant, 7 were in fact discarded (PPV 63%). In comparison, of the 10 kidneys declined by the assessment rule, 9 were discarded (PPV 90%). Of the 83 kidneys described as transplantable by the independent consultants, 69 were transplanted (NPV 83%), whilst of the 84 kidneys described as transplantable by the assessment rule, 72 were transplanted (NPV 86%). See Table 2 [tab 2]. 69 of the 72 kidneys implanted in accordance with the assessment rule survived the first year (96% 1-year graft survival) with a mean eGFR of 45.3ml/min1.73m^2).

We examined the cases where our assessment rule and the independent consultant made different decisions (n=7). Within these, our assessment rule correctly predicted the ultimate fate of the kidney in 6 cases (i.e. 4 implanted and 2 declined). Only one kidney was implanted that would not have been if the rule had been followed. At 1-year post transplant, that kidney survived to follow up and was functioning (eGFR 14ml/min1.73m^2).

DISCUSSION

Proforma design

Neutral terms on rating scales are known to give a psychological benchmark that biases against discrimination, as when a middle option is offered, it is far more likely to be chosen[20,21]. Given that the aim of the appraisal process is to discriminate between organs that are fit for transplantation from those that are not, we omitted neutral categories where possible. Complex characteristics that could not be measured directly e.g. 'mild- or moderate- atherosclerosis' were benchmarked in plain English or given numerical thresholds to minimise subjectivity. This strategy appeared to be effective as ratings using the proforma were highly consistent. For all assessment categories, no two assessments of the same kidney deviated by more than 1-point on the rating scale, suggesting that the descriptors normalised language sufficiently to minimise inter-observer variability.

Assessment scores

First, we examined whether the distribution of scores was consistent with expectations for a cohort of kidneys that were predominantly transplantable. Scores relating to the size of the kidney appeared to be centrally placed and reasonably well-spread across the scale. However, as expected, there was significant skewness in the distributions of the other assessments denoting organ quality, so that most kidneys were scored as normal or only mildly suboptimal. This gives us confidence that the categories are appropriately sized and well placed. Next, we observed whether there was a relationship between the scores given using the assessment proforma and factors which are already known to affect organ quality, most important of which is donor age. We consistently found that whilst low scoring (close to

optimal) kidneys were retrieved from donors of any age, high scoring (e.g. >= 2) kidneys were from a comparatively narrower pool of exclusively older donors. This pattern was observed with respect to the Carrell patch and renal artery scoring, with the corresponding p-values 0.005 and 0.04 by the Kruskal–Wallis tests [fig 6].

No obvious relationship was observed between donor age and Remuzzi score (p-value 0.91) [fig 7], nor were any of the macroscopic assessment scores able to reliably predict the biopsy score. The fact that we were unable to observe these relationships is difficult to interpret. However, we note that the availability - or otherwise - of biopsy scores was not random, since 'good quality' kidneys (based on clinical assessment), as well as kidneys from younger donors (typically aged <60 years) were often simply transplanted and not biopsied. Furthermore, given that fewer than 100 kidneys were included in the study, the very small numbers of highly scoring kidneys likely limited our power to detect a relationship.

Appraisal rule

One aim of the study was to identify factors that reliably predicted organ utilisation within this cohort of donor kidneys. Following our analysis, we found two factors that were important, and we used these to form the basis of an easy to use, rapid organ assessment rule. Within our cohort of kidneys, the rule proposed to identify kidneys at higher risk of decline for transplantation was: "Carrel patch quality + Perfusion quality >3", as it performed well in a retrospective 2x2 contingency table analysis. Importantly, by following this simple rule, assessments made by a mix of junior and senior transplant trainees were more accurate for predicting utilisation than an independent transplant consultant at our centre assessing the same kidney. The choice of an independent consultant at the same centre to compare test

performance was pragmatic, but also served as the strictest possible control, as many confounding factors such as local policy or 'surgical culture' would be shared. Despite this, and in keeping with what others have reported, surgeons – in this case, even from the same centre - do not always agree on whether a given kidney is suitable for transplant, further buttressing the case for an objective method of interrogating these differences of opinion[2,22].

Limitations

This study has some important limitations which we hope to address in future work. A central issue that affects all studies of pre-implantation assessment methods is that there is currently no way of evaluating whether the ultimate decision to discard an organ is the correct one. In our series, we found 9 kidneys were predicted to be at high risk of discard by the CKAT appraisal rule that were ultimately discarded, compared to only 7 by the independent transplant consultant. However, it would be erroneous to assume that this increased sensitivity necessarily implies a better test, as that would depend on the outcomes of the kidneys had they been transplanted. Although this cannot be proved definitively in the absence of functional outcome measures in transplanted patients, further tests of 'transplantability' - like ex-vivo perfusion or pre-implantation biopsy - could be used in combination for kidneys deemed at higher risk of decline following macroscopic assessment, although this would depend on local availability and the outcomes of ongoing trials[6,7,23].

For logistical reasons, we required that surgeons from our centre retrieved, then implanted the kidney, and we were therefore only able to recruit a small number of kidneys over the 2-year study-period, limiting our power to detect potentially important associations. Entry was particularly difficult for DBD kidneys, as they were

nationally allocated at the time of the study, which meant that opportunities for multiple assessments (e.g. by retrieval, and implanting teams) were infrequent. Whilst it is entirely possible that the same macroscopic appearance in a DCD vs DBD kidney could have different implications for organ quality in each case, we feel that the overwhelming evidence from retrospective studies of organ outcome (as well as our own local experience) shows that DCD and DBD kidneys perform equivalently, and by extension should be utilised equivalently in the absence of strong evidence proving a difference[16,24–27]. Moreover, limited numbers meant that we had to develop and test the rule on the same cohort, so our proposed CKAT model risks overfitting our current dataset. Ultimately, these assumptions require testing in a large macroscopic assessment study to address these questions, due to lack of direct evidence in the area. In a future study, higher numbers could also enable us to interrogate the relationship between macroscopic features and biopsy-based assessment, which to our knowledge, has not yet been described.

Future considerations

We found that kidneys that did not meet the CKAT score threshold (Carrel + Perfusion >3) were highly likely to be suitable for transplantation. Therefore, given the high specificity and negative predictive values observed at that threshold, kidneys scoring below the threshold could - in theory - be transplanted with confidence, without the need for additional tests that would incur additional delays and/or cost. However, to confirm this, we would need to validate the rule in a large prospective study, including assessors from a variety of transplant centres. In the UK, retrieval surgeons conform to centrally organised national standards[28]. A future study (justified by this experience) could leverage this national framework by routinely appraising kidneys using the CKAT proforma, whilst simultaneously

collecting outcome data via the UK Transplant Registry, following the template of other low-cost 'registry-based trials'[7]. This would ensure that prior to clinical adoption, the threshold was validated robustly, this would need to include an evaluation of what impact macroscopic selection rules might have on overall kidney utilisation, as this would minimise the risk of a negative impact. Whilst this study provides evidence that macroscopic assessment can be predictive of utilisation, long-term follow up will be essential to observe whether macroscopic appraisal can reliably predict future clinical transplant outcomes, or even whether selection thresholds defined today require future adjustment consequent to the impact of emerging technologies such as biopsy assessment, ex-vivo machine perfusion or medical management.

It is also important to emphasise that kidneys assigned a CKAT-score in excess of the threshold should not be considered 'untransplantable', as the rule performed less well in predicting outcomes above the threshold. Instead, these kidneys are more likely to benefit from more intensive assessment, or further tests of transplantability dependent on local availability. The CKAT assessment tool could therefore be used to complement or rationalise the application of the other assessment techniques. The final decision to proceed with a transplant should always rest with the implanting team, taking into account both donor and recipient factors.

The goal of the CKAT-score is not to supplant clinician decision making, but to begin the process of developing an evidence base on which determinations on organ quality reside. We would argue that, until we have a common, standardised language, the transplant community will continue to struggle to get the right organ to the right recipient. Similarly, until we can be confident that organ descriptors are

being used reliably by different surgeons, we cannot include these terms effectively in our models of organ risk or study their impacts robustly.

In the future, many kidneys will have on-table photographic images taken at retrieval using mobile phones or tablets. Whilst this may reduce the implanting surgeon's reliance on retrieval assessment, it continues the tradition of macroscopic appraisal being subjectively determined, albeit remotely. Future attempts to highlight and limit 'unnecessary discard' of transplants will ultimately rely on a more universally accepted definition of an 'ideal' or 'transplantable' kidney. Whilst known clinical and operative risk factors will certainly play a role, this definition will need to integrate the relevant aspects of macroscopic assessment, as these are part of quality-assessment in clinical practice. Therefore, future work in minimising 'unnecessary discard' is likely to benefit from standardised descriptors of organ quality, potentially informed by the CKAT, or a variant of it.

CONCLUSION

The purpose of this study was to move towards standardised language in the appraisal kidneys for transplantation by producing an objective (minimal inter-observer variability), standardised assessment proforma for surgeons to use whilst assessing kidney transplants (CKAT). We also showed that a simple assessment rule (CKAT-score; Carrel + Perfusion >3) can predict utilisation more accurately than an unblinded, independent consultant transplant surgeon. To our knowledge, this is the first attempt to objectively identify macroscopic features that are relevant to utilisation. Additional benefits of the CKAT appraisal method are that it can be implemented without additional cost, assessment requires minimal prior transplant experience and no additional equipment is required.

TABLES

Table 1A

Donor demographics	Total n=49	Transplanted n=41	Not transplanted n=8
Sex			
Male	35 (71%)	31 (75%)	4 (50%)
Female	14 (29%)	10 (24%)	4 (50%)
Age, (median) years	62	61	69.5
BMI, (median) kg/m^2	26.7	26.1	28.4
Cardiovascular morbidity (HTN/Diabetes)	17 (35%)	13 (32%)	6 (75%)
Terminal creatinine (median) µmol/l	76.5	74.5	95
DCD	44 (95%)	-	-

Table 1B

Assessor demographics	Assessments performed (n=184)
Retrieval assessor's grade	
Junior trainee [a]	84 (46%)
Senior trainee [b]	87 (47%)
Consultant	13 (7%)

a ~3-6 years post-graduation, b >6 years post-graduation

Table summary of donor and assessor demographics.

Table 2

Consultant opinion				
		Predicted by Consultants		Sensitivity & Specification
		Decline	Implant	
Actual Outcomes	Declined	7	14	33%
	Implant	4	69	95%
Positive and Negative Predictive Values		63%	83%	
Assessments made using CKAT + assessment rule				
		Decision by CKAT rule		Sensitivity & Specification
		Decline	Accept	
Actual Outcomes	Declined	9	12	47%
	Accept	1	72	99%
Positive and Negative Predictive Values		90%	86%	

A 2x2 contingency table summarising the predictive performance of an independent transplant consultant compared to actual outcomes. A similar table describes the performance of CKAT assessments using the rule: “Carrel patch quality + Perfusion quality >3” compared to actual outcomes.

FIGURES

Figure 1A

Figure 1 – Part A

SURGEON APPRAISAL of DECEASED DONOR KIDNEYS
Service Evaluation No. 2324

We are conducting an evaluation of surgeon appraisal of deceased donor kidneys to determine whether it is possible to develop an objective tool. We are interested to know what factors relating to the kidney itself are involved in your decision to accept or reject the organ. We would like you to fill out the table overleaf AFTER you have appraised the kidney and documented your decision below (this should be your own evaluation and not that of a senior colleague). If you are sending a biopsy for this kidney prior to making a final decision, please give an assessment decision based on what you would do if you did not have the option of a biopsy.

Date of retrieval: ___ /___ /___ Retrieval Team / Implanting Team (*please circle*)

Organ Donor No: _____ Donor hospital: _____

Name of assessor: _____ Level (Cons/SpR/SHO): _____

Kidney Side (please circle)**:** Left Kidney Right Kidney

Before you fill out the adjacent form, please assess the DCD kidney as you usually would, record your decision to ACCEPT (single or double use) or REJECT it below, and then complete this form.

Based on my own appraisal of the kidney, I would recommend the following for this kidney (please circle):

Accept – use as a *single* transplant **Accept** – use as a *double* transplant **Reject**

Did you consider donor factors in your appraisal of the kidney (please circle): **Yes / No**

If yes, please indicate which factors you considered: **Age, (Pre-)terminal creatinine, Co-morbidities (please list), Other (please list)**

……

Figure 1B

Figure 1 – Part B

For the kidney you have just appraised, now please fill out the form below:

FACTORS SCORES	0	1	2	3
Estimated size of kidney (pole to pole length)	<8cm	8-10cm	10-12cm	>12cm
Carrel (Aortic) patch atherosclerosis	none	mild (patch is still viable)	moderate (patch is still just viable)	heavy (not usable, patch must be cut off)
Renal artery atherosclerosis	none	mild	moderate	heavy (obvious)
Perfusion characteristics	perfect	moderately well (<10% of areas not perfused)	poorly (10-30% of areas not perfused)	very badly perfused (>30% of areas not perfused)
Parenchymal damage – infarction	no evidence of infarction	infarction involving <10% of total kidney surface area	infarction involving 10-30% of total kidney surface area	infarction involving >30% of total kidney surface area
Parenchymal feel/texture/fibrosis	Soft, pliable, no evidence of fibrosis/scarring on kidney surface	Somewhat firm, reduced elasticity ± minimal scarring/fibrosis involving <30% the kidney surface area	Firm, no elasticity, obvious scarring /fibrosis involving >30% of kidney surface area	

TECHNICAL/OTHER DISCARD FACTORS		
Anatomy	Kidney unusable due to aberrant anatomy (not reconstructable):	Yes / No
Retrieval damage	Kidney unusable due to damage during retrieval (not reconstructable):	Yes / No
Malignancy or infection	Kidney unusable due to known or suspected malignancy or infection:	Yes / No

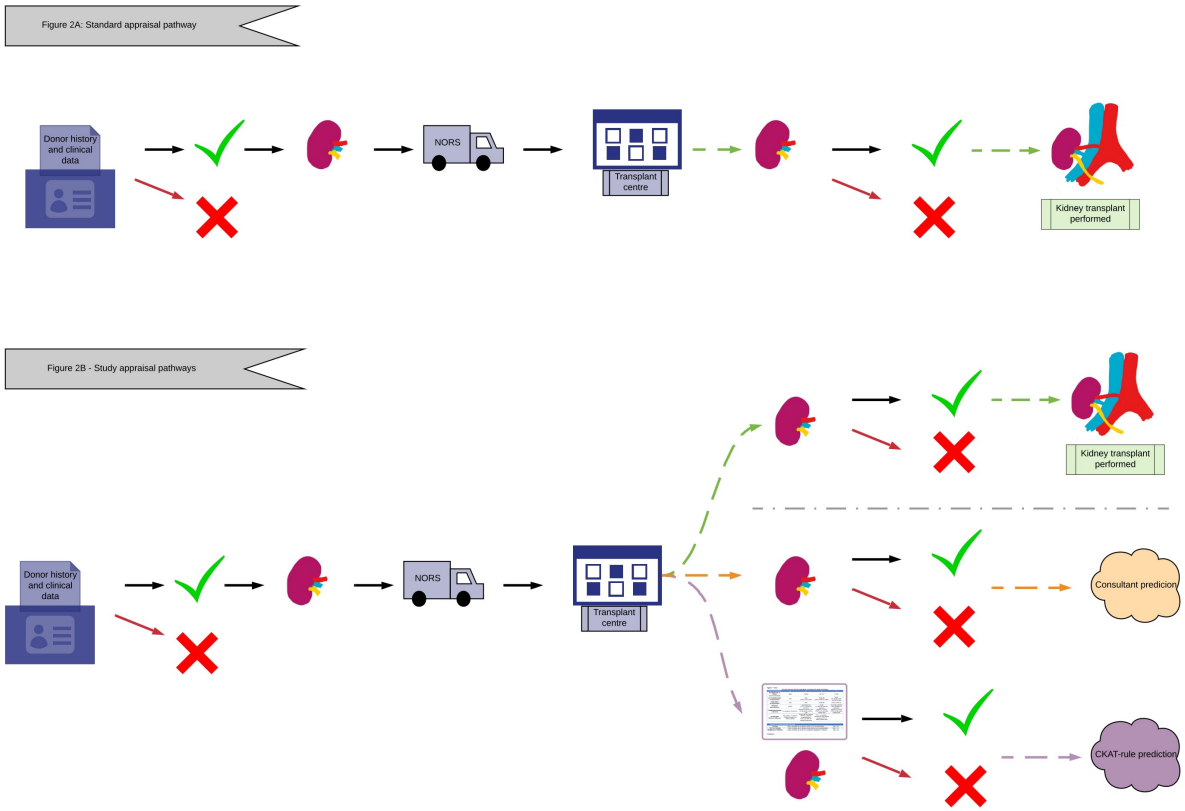
Comments:……

Materials provided to surgeons to record:

A. Their overall assessment of the kidney's suitability for transplantation

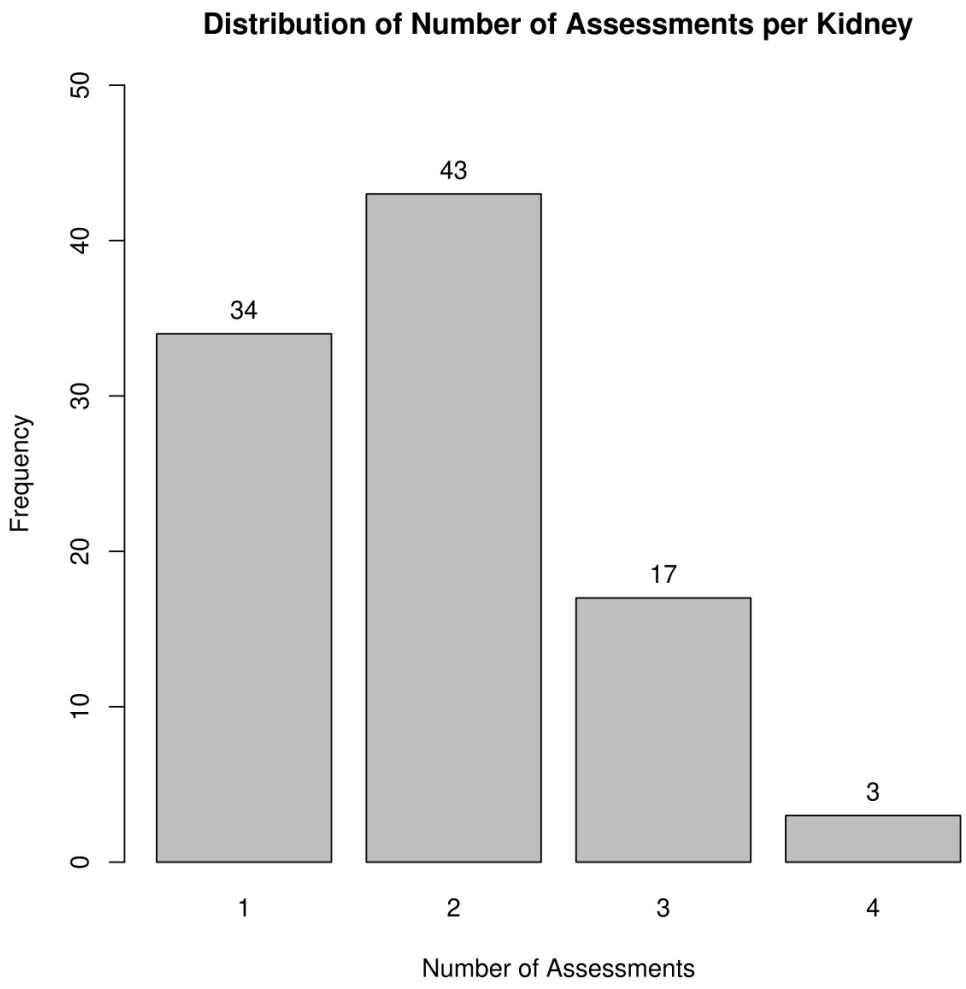
B. Their assessment of the macroscopic appearance of the kidney using the Cambridge Kidney Assessment Tool (CKAT)

Figure 2



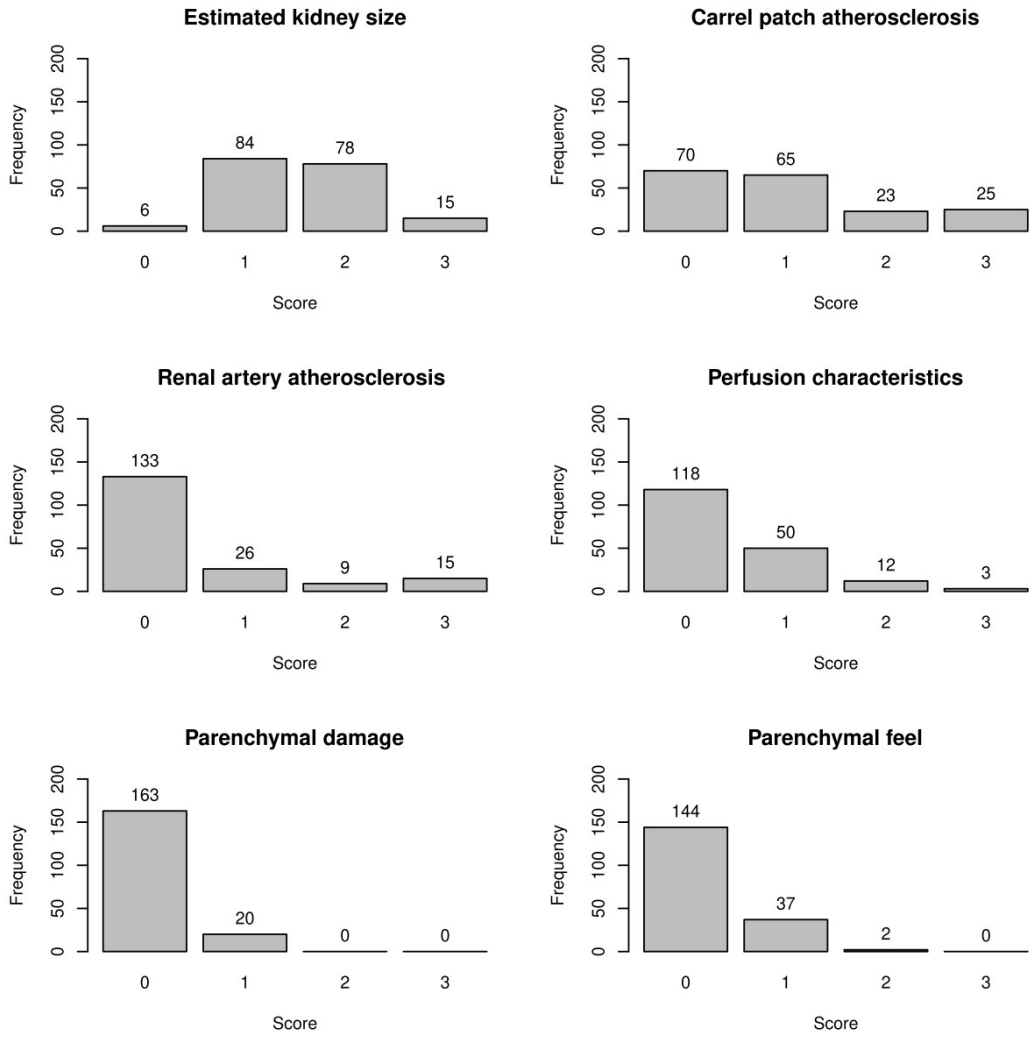
Summary of standard (2A) and study (2B) appraisal pathways. In the standard appraisal pathway, a decision is made to retrieve the kidney based on the donor history and clinical information (including blood tests). These details are confirmed with the NORS team, who retrieve the kidney, assess it macroscopically, and return to the transplant centre with the organ. At the transplant centre, a further macroscopic assessment is made by the implanting surgeon during backbench preparation (green pathway) who then proceeds to implant the kidney or discards it. In the study pathway, two further assessments are made; a second independent consultant makes their own macroscopic assessment (orange pathway), and a transplant surgeon, typically part of the NORS team, uses the CKAT proforma to assess the kidney (purple pathway), the predictive power of these assessments are compared to the actual outcome.

Figure 3



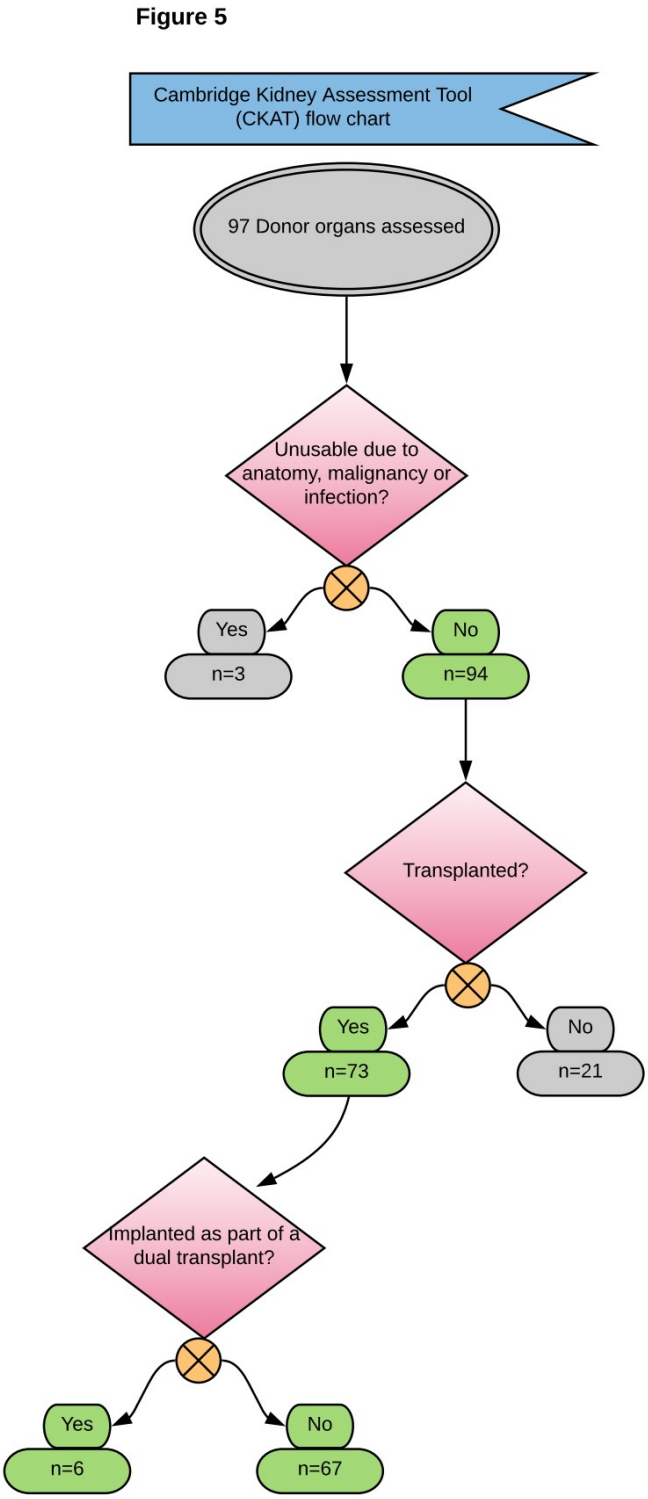
Frequency distribution bar chart showing the number of surgical assessments performed on each kidney in the series, the maximum number of assessments for a single kidney was 4.

Figure 4



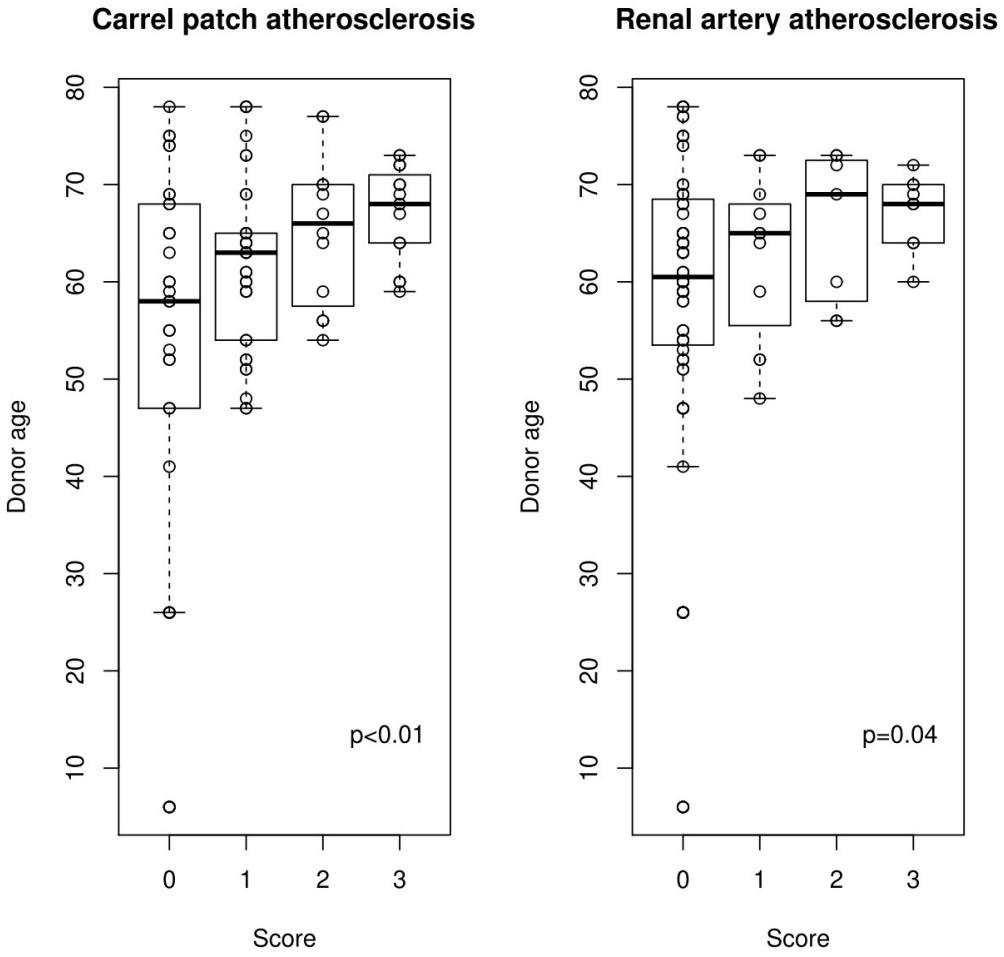
Frequency distribution bar charts summarising the scores given following surgical assessments of kidneys in our series, in each dimension on the CKAT.

Figure 5



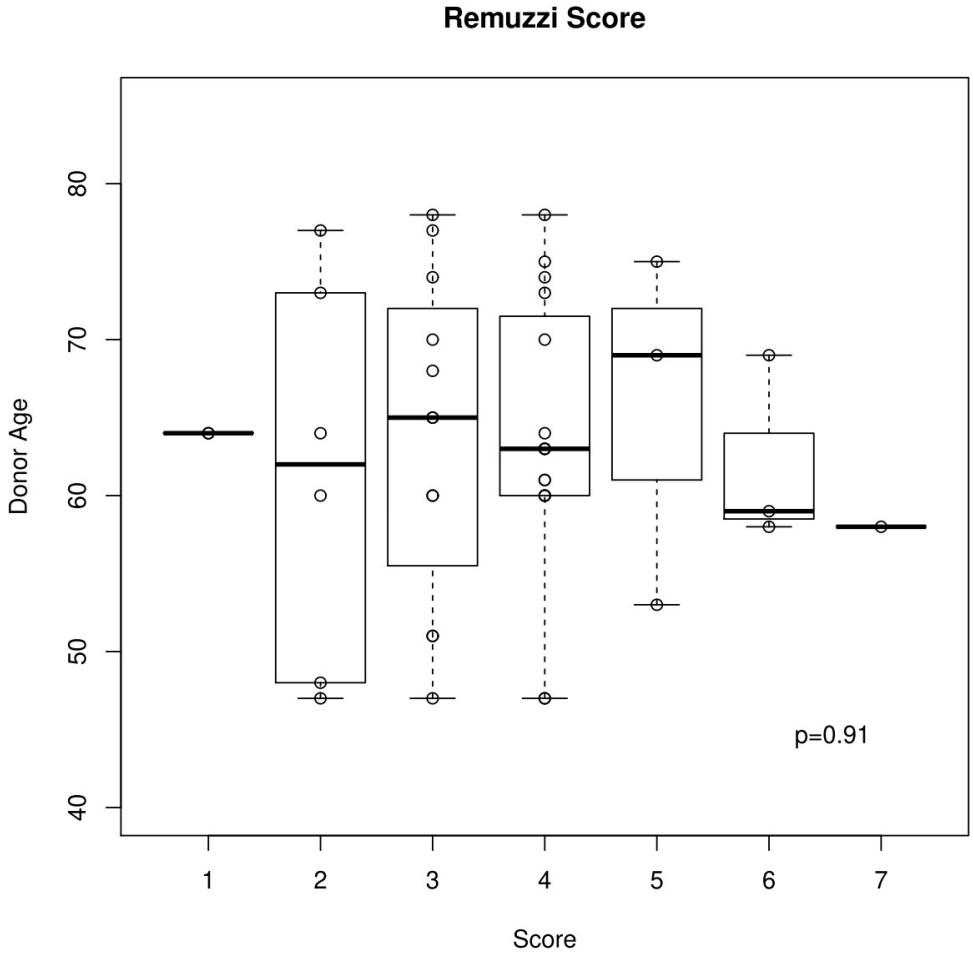
Flow chart summarising the pathway for the kidneys assessed during the study.

Figure 6



Box plots summarising the distribution of donor ages for each category of score given after assessment using the CKAT (Kruskal-Wallis significance test). Higher scores for carrel patch atherosclerosis and renal artery atherosclerosis appear to become restricted to an older cohort of donors, whilst good scores come from donors of any age.

Figure 7



Box plots summarising the distribution of donor ages for each category of Remuzzi score assigned to those kidneys biopsied as part of routine transplant assessment practice at our centre (Kruskal-Wallis significance test).

REFERENCES

1. NHSBT. *Annual Report on Kidney Transplantation*.; 2018. https://nhsbtdbe.blob.core.windows.net/umbraco-assets-corp/12256/nhsbt-kidney-transplantation-annual-report-2017-2018.pdf. Accessed April 28, 2017.

2. Callaghan CJ, Harper SJF, Saeb-Parsy K, et al. The discard of deceased donor kidneys in the UK. *Clin Transpl*. 2014;28(3):345-353. doi:10.1111/ctr.12319

3. Mohan S, Chiles MC, Patzer RE, et al. Factors leading to the discard of deceased donor kidneys in the United States. *Kidney Int*. 2018;94(1):187-198. doi:10.1016/J.KINT.2018.02.016

4. Callaghan CJ, Mumford L, Pankhurst L, Baker RJ, Bradley JA, Watson CJE. Early Outcomes of the New UK Deceased Donor Kidney Fast-Track Offering Scheme. *Transplantation*. 2017;101(12):2888-2897. doi:10.1097/TP.0000000000001860

5. Husain SA, Chiles MC, Lee S, et al. Characteristics and Performance of Unilateral Kidney Transplants from Deceased Donors. *Clin J Am Soc Nephrol*. 2018;13(1):118-127. doi:10.2215/CJN.06550617

6. Hosgood SA, Saeb-Parsy K, Wilson C, Callaghan C, Collett D, Nicholson ML. Protocol of a randomised controlled, open-label trial of ex vivo normothermic perfusion versus static cold storage in donation after circulatory death renal transplantation. *BMJ Open*. 2017;7(1):e012237. doi:10.1136/bmjopen-2016-012237

7. Ayorinde JO, Summers DM, Pankhurst L, et al. PreImplantation Trial of Histopathology In renal Allografts (PITHIA): a stepped-wedge cluster randomised controlled trial protocol. *BMJ Open*. 2019;9(1):e026166. doi:10.1136/BMJOPEN-2018-026166

8. Israni AK, Salkowski N, Gustafson S, et al. New National Allocation Policy for Deceased Donor Kidneys in the United States and Possible Effect on Patient Outcomes. *J Am Soc Nephrol*. 2014;25(8):1842-1848. doi:10.1681/ASN.2013070784

9. Rao PS, Schaubel DE, Guidinger MK, et al. A Comprehensive Risk Quantification Score for Deceased Donor Kidneys: The Kidney Donor Risk Index. *Transplantation*. 2009;88(2):231-236. doi:10.1097/TP.0b013e3181ac620b

10. Dare AJ, Pettigrew GJ, Saeb-Parsy K. Preoperative assessment of the deceased-donor kidney: From macroscopic appearance to molecular biomarkers. *Transplantation*. 2014;97(8):797-807. doi:10.1097/01.TP.0000441361.34103.53

11. Reid AWN, Harper S, Jackson CH, et al. Expansion of the Kidney Donor Pool by Using Cardiac Death Donors with Prolonged Time to Cardiorespiratory Arrest. *Am J Transplant*. 2011;11(5):995-1005. doi:10.1111/j.1600-6143.2011.03474.x

12. Rachel J. Johnson, Susan V. Fuggle, Lisa Mumford, J. Andrew Bradley, John L. R. Forsythe CJR. A New Uk 2006 National Kidney Allocation Scheme for Deceased Heart-beating Donor Kidneys. *Transplantation*. 2010;89(4):387-394.

doi:10.1097/tp.0b013e3181c9029d

13. Kootstra G, Daemen JH, Oomen AP. Categories of non-heart-beating donors. *Transplant Proc*. 1995;27(5):2893-2894. http://www.ncbi.nlm.nih.gov/pubmed/7482956. Accessed November 20, 2018.

14. Wells AC, Rushworth L, Thiru S, et al. Donor kidney disease and transplant outcome for kidneys donated after cardiac death. *Br J Surg*. 2009;96(3):299-304. doi:10.1002/bjs.6485

15. Remuzzi G, Grinyò J, Ruggenenti P, et al. Early experience with dual kidney transplantation in adults using expanded donor criteria. Double Kidney Transplant Group (DKG). *J Am Soc Nephrol*. 1999;10(12):2591-2598.

16. Mallon DH, Riddiough GE, Summers DM, et al. Successful Transplantation of Kidneys From Elderly Circulatory Death Donors by Using Microscopic and Macroscopic Characteristics to Guide Single or Dual Implantation. *Am J Transplant*. 2015;15(11):2931-2939. doi:10.1111/ajt.13349

17. Kosmoliaptsis V, Salji M, Bardsley V, et al. Baseline donor chronic renal injury confers the same transplant survival disadvantage for DCD and DBD kidneys. *Am J Transplant*. 2015;15(3):754-763. doi:10.1111/ajt.13009

18. Chan A-W, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Ann Intern Med*. 2013;158(3):200. doi:10.7326/0003-4819-158-3-201302050-00583

19. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new

prediction equation. Modification of Diet in Renal Disease Study Group. *Ann Intern Med*. 1999;130(6):461-470. http://www.ncbi.nlm.nih.gov/pubmed/10075613. Accessed November 17, 2019.

20. Dassa C, Lambert J. Effects of a neutral answer choice on the reliability and validity of attitude and opinion items. *Can J Progr Eval* . 1997. https://search.proquest.com/docview/1011295151?pq-origsite=gscholar. Accessed November 24, 2018.

21. Moors G. Exploring the effect of a middle response category on response style in attitude measurement. *Qual Quant*. 2008;42(6):779-794. doi:10.1007/s11135-006-9067-x

22. Mittal S, Adamusiak A, Horsfield C, et al. A Re-evaluation of Discarded Deceased Donor Kidneys in the UK: Are Usable Organs Still Being Discarded? *Transplantation*. 2017;101(7):1698-1703. doi:10.1097/TP.0000000000001542

23. Hosgood SA, Saeb-Parsy K, Hamed MO, Nicholson ML. Successful Transplantation of Human Kidneys Deemed Untransplantable but Resuscitated by Ex Vivo Normothermic Machine Perfusion. *Am J Transplant*. 2016;16(11):3282-3285. doi:10.1111/ajt.13906

24. Locke JE, Segev DL, Warren DS, Dominici F, Simpkins CE, Montgomery RA. Outcomes of Kidneys from Donors After Cardiac Death: Implications for Allocation and Preservation. *Am J Transplant*. 2007;7(7):1797-1807. doi:10.1111/j.1600-6143.2007.01852.x

25. Summers DM, Johnson RJ, Allen J, et al. Analysis of factors that affect

outcome after transplantation of kidneys donated after cardiac death in the UK: a cohort study. *Lancet*. 2010;376(9749):1303-1311. doi:10.1016/S0140-6736(10)60827-6

26. Summers DM, Johnson RJ, Hudson A, Collett D, Watson CJ, Bradley JA. Effect of donor age and cold storage time on outcome in recipients of kidneys donated after circulatory death in the UK: a cohort study. *Lancet (London, England)*. 2013;381(9868):727-734. doi:10.1016/S0140-6736(12)61685-7

27. Lisa M. *KAG(18)8 NHS BLOOD AND TRANSPLANT KIDNEY OFFERING SCHEME WORKING GROUP ENDORSEMENT OF A NEW NATIONAL KIDNEY OFFERING SCHEME*. https://youtu.be/iOyeGPh85K0. Accessed August 30, 2019.

28. Zalewska K, Ploeg R. National standards for organ retrieval from deceased donors. Version 2.6. *NHS Blood Transpl*. 2013;MPD1043/8:1-50. https://nhsbtdbe.blob.core.windows.net/umbraco-assets-corp/12548/mpd1043-nors-standard.pdf.