

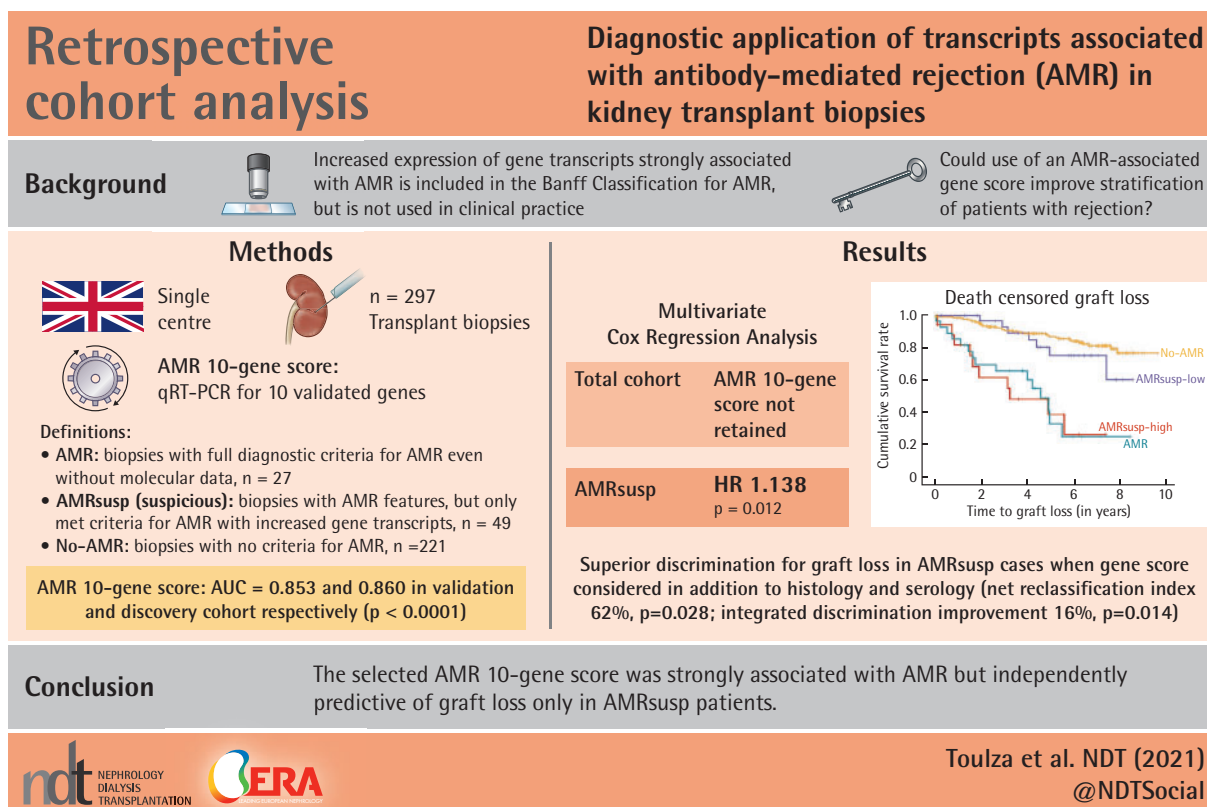
# Diagnostic application of transcripts associated with antibody-mediated rejection in kidney transplant biopsies

Frederic Toulza <sup>1</sup>, Kathy Dominy<sup>2</sup>, Michelle Willicombe<sup>1,3</sup>, Jack Beadle<sup>1</sup>, Eva Santos<sup>4</sup>, H. Terence Cook<sup>1</sup>, Richard M. Szydlo<sup>5</sup>, Adam McLean<sup>3</sup> and Candice Roufousse<sup>1</sup>

<sup>1</sup>Department of Immunology and Inflammation, Imperial College, Centre for Inflammatory Disease, London, UK; <sup>2</sup>Molecular Pathology Laboratory, North West London Pathology, London, UK; <sup>3</sup>Imperial College Renal and Transplant Centre, Imperial College Healthcare NHS Trust, London, UK; <sup>4</sup>Histocompatibility and Immunogenetics, North West London Pathology, London, UK; <sup>5</sup>Department of Immunology and Inflammation, Imperial College, London, UK

Correspondence to: Candice Roufousse; E-mail: c.roufousse@imperial.ac.uk

## GRAPHICAL ABSTRACT



## ABSTRACT

**Background.** The diagnosis of antibody-mediated rejection (AMR) is reached using the Banff Classification for Allograft

Pathology, which now includes gene expression analysis. In this study, we investigate the application of ‘increased

expression of thoroughly validated gene transcripts/classifiers strongly associated with AMR' as diagnostic criteria.

**Method.** We used quantitative real-time polymerase chain reaction for 10 genes associated with AMR in a retrospective cohort of 297 transplant biopsies, including biopsies that met the full diagnostic criteria for AMR, even without molecular data (AMR,  $n = 27$ ), biopsies that showed features of AMR, but that would only meet criteria for AMR with increased transcripts [suspicious for AMR (AMRsusp),  $n = 49$ ] and biopsies that would never meet criteria for AMR (No-AMR,  $n = 221$ ).

**Results.** A 10-gene AMR score trained by a receiver-operating characteristic to identify AMR found 16 cases with a high score among the AMRsusp cases (AMRsusp-high) that had significantly worse graft survival than those with a low score (AMRsusp-low;  $n = 33$ ). In both univariate and multivariate Cox regression analysis, the AMR 10-gene score was significantly associated with an increased hazard ratio (HR) for graft loss (GL) in the AMRsusp group (HR = 1.109,  $P = 0.004$  and HR = 1.138,  $P = 0.012$ , respectively), but not in the whole cohort. Net reclassification index and integrated discrimination improvement analyses demonstrated improved risk classification and superior discrimination, respectively, for GL when considering the gene score in addition to histological and serological data, but only in the AMRsusp group, not the whole cohort.

**Conclusions.** This study provides evidence that a gene score strongly associated with AMR helps identify cases at higher risk of GL in biopsies that are suspicious for AMR but do not meet full criteria.

**Keywords:** acute rejection, chronic renal failure, gene expression, graft failure, kidney biopsy

## INTRODUCTION

Antibody-mediated rejection (AMR) is the most common cause of allograft rejection and loss, and occurs when circulating donor-specific antibodies (DSAs) interact with the graft endothelium, triggering endothelial cell activation and recruitment of inflammatory cells to the vasculature via a cascade of both complement-dependent and -independent mechanisms. This can cause acute graft injury and lead to chronic graft injury through laying down of matrix in glomeruli, peritubular capillaries and arteries. The diagnosis of AMR is reached using the Banff Classification for Allograft Pathology and requires three elements: (i) histological features of AMR (Criterion 1); (ii) evidence of interaction between a DSA and the endothelium (Criterion 2); and (iii) evidence of a DSA (Criterion 3). Following the discovery of increased endothelial and natural killer (NK)-cell-related transcripts in biopsies with AMR [1–4], the Banff Classification was modified in 2013 and 2017, to include 'increased expression of thoroughly validated gene transcripts/classifiers in the biopsy tissue strongly associated with AMR' either as evidence of interaction of a DSA with tissue (Criterion 2) and/or as evidence of a DSA itself (Criterion 3) [5]. Therefore, in the presence of histological features of AMR (Criterion 1), 'increased expression of thoroughly validated

gene transcripts/classifiers in the biopsy tissue strongly associated with AMR' defines cases with AMR.

Several studies have confirmed increased expression of endothelial-associated or NK-associated transcripts in AMR, using microarray analysis [6, 7], quantitative real-time polymerase chain reaction (qRT-PCR) [8] or Nanostring nCounter analysis [9, 10]. Using microarray analysis, Sellares *et al.* showed that the AMR gene signature was associated with increased graft loss (GL) [11, 12], and Loupy *et al.* showed that adding an AMR molecular score to serology and histology data improved stratification of patients with AMR for their risk of GL [12]. We have previously reported that in patients with a *de novo* DSA, a subset of transcripts associated with AMR measured using qRT-PCR is associated with GL [8]. None of these studies, however, investigates the precise application of molecular analysis in the diagnosis of AMR as defined in the Banff Classification.

The aim of this study was to investigate an AMR-associated gene score in the context of use proposed in the Banff classification, by analysing whether increased expression of the AMR-associated gene set defines, amongst cases meeting Criterion 1 alone, a group of cases more closely aligned with AMR in terms of their outcome.

## MATERIALS AND METHODS

### Study population and data collection

The study population comprised patients from Imperial College Healthcare National Health Service (NHS) Trust with a transplant biopsy between 20 July 2010 and 6 November 2018, for which a sample in RNAlater was obtained with sufficient RNA available (500 ng). Renal transplant biopsy cores were obtained under ultrasound guidance with an 18-gauge spring-loaded needle. At least half or a full core was placed in RNAlater (Life Technologies, Paisley, UK). Tissue was obtained from the Imperial College Healthcare NHS Trust Tissue Bank, which has ethics approval to both collect human tissue and release material to researchers (MREC 17/WA/0161). Patients were followed up to February 2020.

### Clinical and pathological features

Cases were classified according to Banff 2017 [5] criteria, using histological and serological data, excluding results of gene expression analysis. Acute tubular injury on its own was not considered a criterion for histological evidence of AMR. Cases identified in this way as meeting the full diagnostic criteria for AMR according to Banff 2017 are designated 'AMR'.

Cases that fulfilled Criterion 1 but that would only meet a diagnosis of AMR with 'increased expression of thoroughly validated gene transcripts/classifiers strongly associated with AMR' were designated suspicious for AMR (AMRsusp).

Criterion 1 was considered fulfilled in cases with  $g+$  peritubular capillaritis (ptc)  $>0$  [with the usual restrictions related to presence of glomerulonephritis or borderline/T-cell-mediated rejection for glomerulitis (g) and ptc, respectively] and/or  $v$  and/or thrombotic microangiopathy (TMA) and/or glomerular capillary wall double contours (cg)  $>1a$  or ptcbml (defined

according to the current Banff definition of AMR [13]). Borderline for T-cell-mediated rejection was defined as at least i1 and at least t1. No-AMR designates the cases that did not fulfil Criterion 1.

### Serological detection of anti-HLA DSAs

DSAs were assessed using LABScreen mixed beads (One Lambda, Inc., Canoga Park, CA, USA) and, if positive, the anti-HLA antibody specificity was identified using LABScreen single antigen beads as previously described [8].

### RNA extraction, qPCR and Z-score transformation

RNA was extracted from the whole tissue as previously described [8]. RNA concentration and purity were quantified with a NanoDrop 2000c Spectrophotometer (LabTech, East Sussex, UK). Total RNA was converted directly into cDNA with random hexamer priming [8]. qRT-PCR was carried out using an Applied Biosystems Vii7 real-time qPCR machine [8]. Gene-specific primers spanning an intron were designed for 18 genes from a list of 'top hits' for AMR in the literature (Supplementary data, Table S1) [1–3, 5, 8, 10, 14, 15]. Gene expression levels were normalized to housekeeping gene *HPRT1* and results were measured relative to Stratagene qRT-PCR Reference RNA (Agilent Technologies using the  $\Delta\Delta C_t$  method) [16].

To compare gene expression levels between genes with differing levels of expression, we performed Z-score normalization [ $Z = (x - \text{mean})/SD$ ] across all samples tested.

### Data analysis

Descriptive data are presented in tables and were analysed using chi-square for categorical parameters and ANOVA or Kruskal–Wallis (KW) for continuous parameters. Correlation between parameters was analysed using Spearman's rank correlation and illustrated using heat map analysis with hierarchical clustering by Euclidean distance (Cluster 3.0 [17]). P-values were corrected using Bonferroni correction. Groups were compared using the Mann–Whitney test. Gene cut-off was determined using receiver operating characteristic (ROC) curve analysis. Outcome analysis was performed using Kaplan–Meier survival curves and Cox proportional hazards regression.

Analyses were performed in IBM SPSS version 25 (IBM Corp., Chicago, IL, USA) and on R version 4.0.3.

## RESULTS

### Study population, demographic and histological characteristics

From an initial set of 665 archived cDNA samples [8], 365 samples had sufficient cDNA (>500 ng) for the current analysis. Forty-three samples were excluded due to technical problems during the PCR analysis (poor triplicates and gene expression levels for housekeeping genes with ct value >34); another 25 were excluded because they were from the same patient. Where several biopsies from the same patient were taken, only one biopsy was included, either the first sample with rejection or, if no biopsies showed rejection, the first chronological

sample. The total number of samples retained for analysis was 297, comprising 27 AMR, 49 AMRsusp and 221 No-AMR.

Demographic characteristics are presented in Table 1. In comparison with the other groups, the AMR group was younger ( $P = 0.007$ ) and had a higher proportion of females ( $P = 0.0003$ ).

Histological features are presented in Table 2. There were higher scores for g, ptc, intimal arteritis (v) and cg, arteriolar hyalinosis (ah) and microcirculation inflammation (MI) in the AMR and AMRsusp groups compared with No-AMR. Details of the main diagnoses in each group are presented in Supplementary data, Table S2.

Of the 49 AMRsusp cases, 15 were included for TMA. Thirty-seven cases were included because of the presence of  $g + ptc$  MI > 0. Of those, 23 had an MI = 1 (2/23 also with features of chronicity), 17/23 with isolated g and 6/23 with isolated ptc. Most of these MI = 1 cases did not have complement component 4 (C4d; 21/23) or evidence of a DSA (22/23). The remaining 14 cases had an MI > 1 (10/14 also with features of chronicity), without C4d or DSA (Table 2).

### Selection of the AMR-associated gene panel and testing of its association with AMR

'Top hits' for AMR from the literature were selected. We discounted genes with >50% missing values in the qRT-PCRs, due to a ct value >34 and/or bad triplicates. We tested each gene for its association with AMR in a discovery ( $n = 151$ ) and a validation cohort ( $n = 146$ ), using a Mann–Whitney test and ROC curve analysis. We retained genes with significant association with AMR in both tests and both cohorts (Supplementary data, Figure S1): *PECAM-1*, *GNLY*, *DARC*, *MYBL1*, *CXCL11*, *KLRF1*, *KLF4*, *CXCL10*, *PLA1a* and *SH2d1b*. We derived an AMR 10-gene score using the sum of Z-scores of the 10 retained genes. The 10-gene score predicted a diagnosis of AMR in both discovery and validation cohorts (Supplementary data, Figure S1).

The AMR 10-gene score was significantly higher in AMR than in AMRsusp ( $P < 0.001$ ) and No-AMR ( $P < 0.001$ ; Figure 1A). Expression in the AMRsusp group was also significantly higher than the No-AMR group ( $P < 0.001$ ). Within the AMR group, there was no significant difference in the AMR 10-gene score between cases with features of AMR chronicity and those without (data not shown). The AMR 10-gene score was significantly different between the AMR group and TCMR cases in the No-AMR group (Mann–Whitney,  $P = 0.0001$ ). There were 16 ABO-incompatible (ABOi) transplants in the No-AMR group; there was no significant difference of expression between AB Oi and ABO compatible cases (Mann–Whitney,  $P = 0.836$ ).

We tested the correlation between the AMR 10-gene score and histological and serological features, using Spearman rank correlation (Figure 1B) as previously described [18]. The AMR 10-gene score showed significant correlation with all histological features of rejection [g, ptc, tubulitis (t), interstitial (non-scarring) inflammation (i), v and cg], and also with features of chronicity.

**Table 1. Demographic and biopsy data**

Characteristic	Study cohort (n = 297)	AMR (n = 27)	AMRsusp (n = 49)	No-AMR (n = 221)	P-value	Test type
<b>Recipients</b>						
Age (at transplant), mean ± SD, years	45.9 ± 13.45	38.3 ± 12.0	43.7 ± 15.7	47.41 ± 12.76	<b>0.007</b>	ANOVA
Sex (female), n (%)	109 (36.7)	17 (63.0)	22 (44.9)	70 (31.7)	<b>0.003</b>	Chi-square
<b>Ethnicity, n (%)</b>						
Caucasian	162 (54.5)	19 (70.4)	22 (44.9)	121 (54.8)	0.101	Chi-square
Asian	82 (27.6)	4 (14.8)	13 (26.5)	65 (29.4)	0.273	Chi-square
Black	34 (11.4)	3 (11.1)	8 (16.3)	23 (10.4)	0.499	Chi-square
Others	18 (6.1)	1 (3.7)	8 (10.2)	12 (5.4)	<b>0.020</b>	Chi-square
Repeat transplantation, n (%)	8 (2.7)	0 (0)	2 (4.1)	6 (2.7)	0.574	Chi-square
<b>Data at transplantation</b>						
<b>Type of transplantation</b>						
ABOi, n (%)	20 (6.7)	1 (3.7)	3 (6.1)	16 (7.2)	0.773	Chi-square
SPK, n (%)	13 (4.4)	1 (3.7)	3 (6.1)	9 (4.1)	0.804	Chi-square
LD, n (%)	126 (42.4)	14 (51.9)	18 (36.7)	94 (42.5)	0.442	Chi-square
DD, n (%)	134 (45.1)	10 (37)	25 (51.0)	99 (44.8)	0.494	Chi-square
HLA-A/B/DR mismatches, mean ± SD	3.24 ± 1.55	3.27 ± 1.59	2.98 ± 1.64	3.29 ± 1.53	0.217	ANOVA
<b>DSAs at biopsy, n (%)</b>						
Class I, n (%)	11 (3.7)	3 (11.1)	2 (4.1)	6 (2.7)	<b>&lt;0.001</b>	Chi-square
Class II, n (%)	18 (6.0)	14 (51.8)	0 (0)	4 (1.8)	<b>&lt;0.001</b>	Chi-square
Both, n (%)	3 (1.0)	2 (7.4)	0 (0)	1 (0.5)	<b>0.002</b>	Chi-square
<b>Follow-up</b>						
Follow-up post-biopsy, median (IQR), years	4.92 (4.69)	3.74 (3.49)	4.34 (4.25)	5.95(4.75)	0.001	KW
<b>Biopsy data</b>						
<b>Time post-transplant to biopsy</b>						
Indication, median (IQR), years	3.7 (4.9)	4.9 (4.7)	4.6 (5.7)	3.4 (4.2)	0.099	KW
Surveillance, median (IQR), years	2.3 (2.9)	8.7 (NA <sup>a</sup> )	2.2 (2.0)	2.1 (3.0)	0.907	KW
<b>Individual histological Banff scores</b>						
Glomerular number, mean ± SD	16.8 ± 7.5	16.8 ± 8.3	16.9 ± 7.6	16.7 ± 7.4	0.983	ANOVA
IFTA nearest 10%, % mean ± SD	15.8 ± 16.9	18.3 ± 17.3	19.9 ± 21.2	14.7 ± 15.8	0.1	ANOVA
t, mean ± SD	0.45 ± 0.81	0.78 ± 1.12	0.45 ± 0.82	0.41 ± 0.76	0.085	ANOVA
i, mean ± SD	0.28 ± 0.74	0.59 ± 1.12	0.31 ± 0.79	0.23 ± 0.66	0.052	ANOVA
ct, mean ± SD	0.02 ± 0.16	0.11 ± 0.32	0.06 ± 0.32	0.00 ± 0.00	<b>0.001</b>	ANOVA
ci, mean ± SD	0.75 ± 0.91	1.07 ± 1.07	0.88 ± 0.92	0.68 ± 0.89	0.060	ANOVA
ct, mean ± SD	0.96 ± 0.84	1.11 ± 0.89	1.04 ± 0.83	0.93 ± 0.83	0.439	ANOVA
ci, mean ± SD	0.93 ± 0.86	1.11 ± 0.89	1.04 ± 0.88	0.89 ± 0.85	0.316	ANOVA
cv, mean ± SD	1.17 ± 0.86	1.19 ± 0.78	1.1 ± 0.80	1.19 ± 0.88	0.825	ANOVA
g, mean ± SD	0.27 ± 0.65	1 ± 1.08	0.82 ± 0.95	0.06 ± 0.28	<b>&lt;0.001</b>	ANOVA
ptc, mean ± SD	0.26 ± 0.64	1.44 ± 0.85	0.51 ± 0.77	0.06 ± 0.33	<b>&lt;0.001</b>	ANOVA
MI, mean ± SD	0.54 ± 1.08	2.44 ± 1.50	1.33 ± 1.28	0.12 ± 0.44	<b>&lt;0.001</b>	ANOVA
C4d score mean ± SD	0.64 ± 0.85	1.56 ± 1.09	0.38 ± 0.64	0.59 ± 0.79	<b>&lt;0.001</b>	ANOVA
cg, mean ± SD	0.19 ± 0.62	0.93 ± 1.11	0.50 ± 1.03	0.02 ± 0.15	<b>&lt;0.001</b>	ANOVA
ah, mean ± SD	0.7 ± 0.92	1.07 ± 0.99	0.9 ± 0.98	0.61 ± 0.88	<b>0.012</b>	ANOVA

Chi-square test was applied for categorical data, and ANOVA or Kruskal Wallis (KW) for continuous data. ABOi, AB blood group incompatible; SPK, simultaneous pancreas/kidney transplant; LD, Live Donor; DD, Deceased Donor; IFTA, Interstitial fibrosis/tubular atrophy; Banff lesions scores: t, tubulitis; i, interstitial(non-scarred) inflammation; v, intimal arteritis; ti, total inflammation; ct, tubular atrophy; ci, interstitial fibrosis; cv, vascular fibrous intimal; g, glomerulitis; ptc, peritubularcapillaritis; MI, microcirculation inflammation; cg, glomerular capillary wall doublecontours; ah, arteriolar hyalinosis. IQR, interquartile range. Bold values, denote statistical significance at the  $P < 0.05$ .

<sup>a</sup>NA, not applicable; this is due to low number of events.

### Graft outcome analysis without and with molecular data

There were 65 allograft failures, 15 in the AMR group, 17 in the AMRsusp group and 33 in the no-AMR group. Death-censored Kaplan–Meier survival curve for GL from time of biopsy was carried out comparing AMR, AMRsusp and No-AMR (Figure 2A). There was significantly worse graft survival for AMR compared with AMRsusp ( $P = 0.047$ ) and No-AMR ( $P < 0.0001$ ). AMRsusp had an intermediate graft survival rate between No-AMR and AMR. The survival at 5 years was 89% for No-AMR, 69% for AMRsusp and 48% for AMR group.

To determine a cut-off value for the AMR 10-gene score, we performed ROC curve analysis on the full cohort (Figure 2B). The area under the curve (AUC) for a diagnosis of AMR versus No-AMR was 0.838 [95% confidence interval (CI) 0.749–0.943,  $P < 0.0001$ ]. Youden’s J statistic identified a cut-off value of 0.23, which was used to split AMRsusp samples, by gene score, into AMRsusp-high ( $n = 16/49$ ) and AMRsusp-low ( $n = 33/49$ ). The 5-year survival was 44% for AMRsusp-high and 82% for AMRsusp-low.

The characteristics of the AMRsusp-high and AMRsusp-low groups are compared in Table 2. AMRsusp-high showed

**Table 2. Comparison of histology and serology between AMRsusp-low and AMRsusp-high**

Characteristic	AMRsusp total (n = 49)	AMRsusp-low (n = 33)	AMRsusp-high (n = 16)	P-value	Test type
Post-transplant time to biopsy, median (IQR), years	1.65 (4.19)	1.35 (2.89)	3.62 (6.05)	0.060	KW
DSA positive, n (%)	2 (4.1)	1 (3.0)	1 (6.2)	0.593	Chi-square
C4d positive, n (%)	2 (4.1)	1 (3.0)	1 (6.2)	0.593	Chi-square
AMR chronicity features, n (%)	14 (28.6)	10 (30.3)	4 (25)	0.699	Chi-square
Criteria 1, n (%)					
TMA	15 (30.6)	10 (30.3)	5 (31.2)	0.946	Chi-square
V	2 (4.1)	1 (3.0)	1 (6.2)	0.593	Chi-square
MI = 1	23 (46.9)	17 (51.5)	6 (37.5)	0.356	Chi-square
MI > 1	14 (28.6)	7 (21.2)	7 (42.7)	0.101	Chi-square
IFTA nearest 10%, mean ± SD	19.98 ± 21.18	13.64 ± 10.91	33.06 ± 30.18	<b>0.002</b>	ANOVA
Banff lesion scores, mean ± SD					
t	0.45 ± 0.82	0.39 ± 0.70	0.56 ± 1.03	0.765	ANOVA
i	0.31 ± 0.79	0.18 ± 0.46	0.56 ± 1.21	0.117	ANOVA
v	0.06 ± 0.32	0.03 ± 0.17	0.13 ± 0.5	0.332	ANOVA
ti	0.88 ± 0.93	0.7 ± 0.73	1.25 ± 1.18	<b>0.021</b>	ANOVA
ct	1.04 ± 0.86	0.85 ± 0.71	1.44 ± 1.03	<b>0.024</b>	ANOVA
cv	1.1 ± 0.80	1.06 ± 0.78	1.19 ± 0.83	0.371	ANOVA
g	0.82 ± 0.95	0.85 ± 0.97	0.75 ± 0.93	0.276	ANOVA
ptc	0.51 ± 0.77	0.33 ± 0.64	0.87 ± 0.88	<b>0.019</b>	ANOVA
MI	1.33 ± 1.28	1.18 ± 1.26	1.62 ± 1.31	0.515	ANOVA
C4d	0.38 ± 0.64	0.31 ± 0.54	0.50 ± 0.82	0.344	ANOVA
cg	0.5 ± 1.03	0.56 ± 1.05	0.37 ± 1.02	0.558	ANOVA
ah	0.9 ± 0.98	0.79 ± 0.96	1.12 ± 1.02	0.265	ANOVA

Chi-square test was applied for categorical data, and ANOVA test for continuous data. DSA, Donor Specific Antibody; C4d, Complement component 4, AMR, antibody mediated rejection; TMA, Thrombotic Microangiopathy; v, intimal arteritis; MI, microcirculation inflammation; IFTA, Interstitial fibrosis/tubular atrophy; IQR, interquartile range; t, tubulitis; i, interstitial (non-scarred) inflammation; ti, total inflammation; ct, tubular atrophy; cv, vascular fibrous intimal; g, glomerulitis; ptc, peritubular capillaritis; cg, glomerular capillary wall double contours; ah, arteriolar hyalinosis. Bold values, denote statistical significance at the  $P < 0.05$ .

significantly more interstitial fibrosis/tubular atrophy (ct), a higher total inflammation (ti) score, a higher ptc score and a non-significant trend to being performed later in time post-transplant.

Death-censored Kaplan–Meier survival curve for GL from time of biopsy was performed comparing AMRsusp-low and AMRsusp-high (Figure 2C). We observed a significantly better survival rate in AMRsusp-low compared with AMRsusp-high ( $P < 0.0001$ ). Survival in the AMRsusp-high aligned with the AMR group ( $P = 0.335$ ), and AMRsusp-low aligned with the No-AMR group ( $P = 0.563$ ).

Applying molecular data to the whole cohort, there was a significant difference in survival curves between cases with high and low AMR 10-gene scores ( $P < 0.001$ ) (Supplementary data, Figure S2A). However, there was no significant difference in survival curves between No-AMR-low and No-AMR-high ( $P = 0.431$ ; Supplementary data, Figure S2B) or between AMR-low and AMR-high ( $P = 0.889$ ; Supplementary data, Figure S2C).

### Uni- and multivariate analysis of serological, histological and molecular features for predicting kidney graft failure

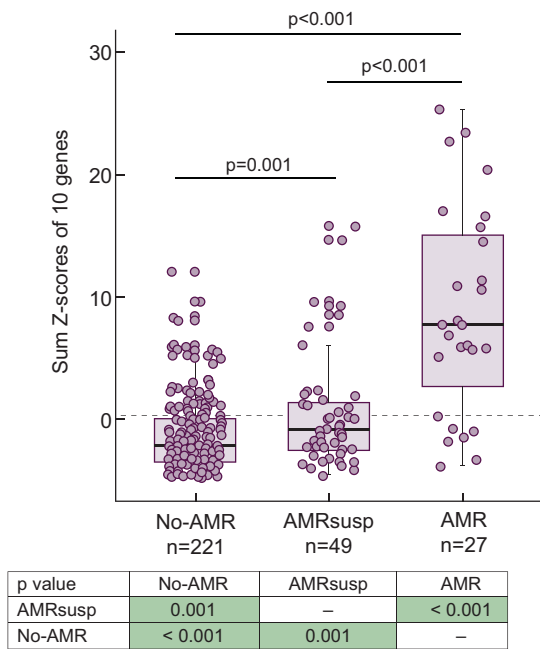
We used Cox proportional hazard regression to predict graft failure, in both uni- and multivariate analyses, in the whole cohort of biopsies (Table 3). Significant predictors for GL in the univariate analysis included type of biopsy (indication versus surveillance) [hazard ratio (HR) = 2.783,  $P = 0.007$ ], time from transplant to biopsy (HR = 1.060,  $P = 0.0007$ ), the presence of a class II DSA (HR = 3.014,  $P \leq 0.001$ ), a diagnosis of AMR

(HR = 6.012,  $P < 0.001$ ) or AMRsusp (HR = 2.833,  $P < 0.001$ ), an MI score of 1 (HR = 2.358,  $P = 0.011$ ) or MI > 1 (HR = 5.626,  $P < 0.001$ ),  $v > 0$  (HR = 3.923,  $P = 0.021$ ), IFTA% (HR = 1.030,  $P < 0.001$ ),  $cg > 0$  (HR = 4.201,  $P < 0.001$ ), and the AMR 10-gene score as a dichotomous (HR = 2.389,  $P = 0.001$ ) or a continuous (HR = 1.091,  $P < 0.001$ ) parameter. In the multivariate analysis, only type of biopsy (HR = 2.492,  $P = 0.028$ ), MI > 1 (HR = 2.543,  $P = 0.042$ ) and IFTA% (HR = 1.029,  $P < 0.001$ ) were retained.

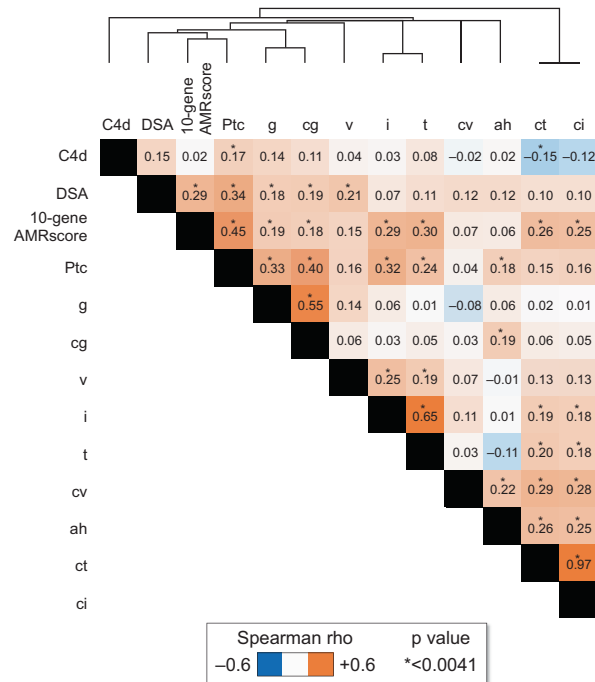
The same analysis was applied to the AMRsusp group (Table 4). In the univariate analysis, time from transplant to biopsy (HR = 1.069,  $P = 0.025$ ), IFTA% (HR = 1.034,  $P = 0.001$ ), and the AMR 10-gene score either dichotomous (HR = 3.844,  $P = 0.008$ ) or continuous (HR = 1.109,  $P = 0.004$ ) were significant. In the multivariate analysis, only IFTA% (HR = 1.032,  $P = 0.022$ ) and the continuous AMR 10-gene score (HR = 1.138,  $P = 0.012$ ) were retained.

Finally, to simulate the clinical use of these variables, logistic regression models were generated using different combinations of three variables: (H) for histology of AMR (as described in Materials and methods), (D) for DSA and/or C4d positivity and (G) for an AMR 10-gene score above threshold (Supplementary data, Figure S3; according to previously published method [18]). The full D + H + G model incorporating all variables demonstrated a larger ROC AUC than each of the reduced models (AUC = 70.6% in the full cohort, 69.3% in the AMRsusp group) although this did not reach statistical significance when compared with D + H in either group. Net reclassification index (NRI) analysis demonstrated improved risk classification with the full D + H +

**A Expression of AMR 10-gene score in AMR, AMR susp and No-AMR groups**



**B Correlation between histological, serological and gene expression parameters**



**FIGURE 1:** Expression and correlations of the AMR 10-gene score. (A) Expression of AMR 10-gene score in AMR, AMRsusp and No-AMR groups. The box plots represent the distribution of the AMR 10-gene score expression in three groups: AMR ( $n = 27$ ), AMRsusp ( $n = 49$ ) and No-AMR ( $n = 221$ ). Each dot represents one individual sample. Mann–Whitney statistical test was applied comparing groups two by two, with P-values in the table below. (B) Correlation between histological, serological and gene expression parameters. The heatmap represents correlation between parameters. The colour intensity in each box represents the level of correlation measure (Spearman rho) between the conditions, with orange representing positive correlation and blue negative correlation. The asterisks in the boxes relate to significance corrected using Bonferroni method ( $*P < 0.0041$ ). Parameters are ordered using a hierarchical clustering by Euclidean distance according to the Spearman rho value.

G model versus the D + H reduced model in the AMRsusp group (NRI = 62.1%,  $P = 0.028$ ), but not in the full cohort (NRI = 26.1%,  $P = 0.059$ ). Integrated discrimination improvement (IDI) also demonstrated superior discrimination with the D + H + G model compared with D + H only in the AMRsusp group (IDI = 16.2%,  $P = 0.014$ ) but not in the full cohort (IDI = 1.6%,  $P = 0.112$ ) (Supplementary data, Figure S3B).

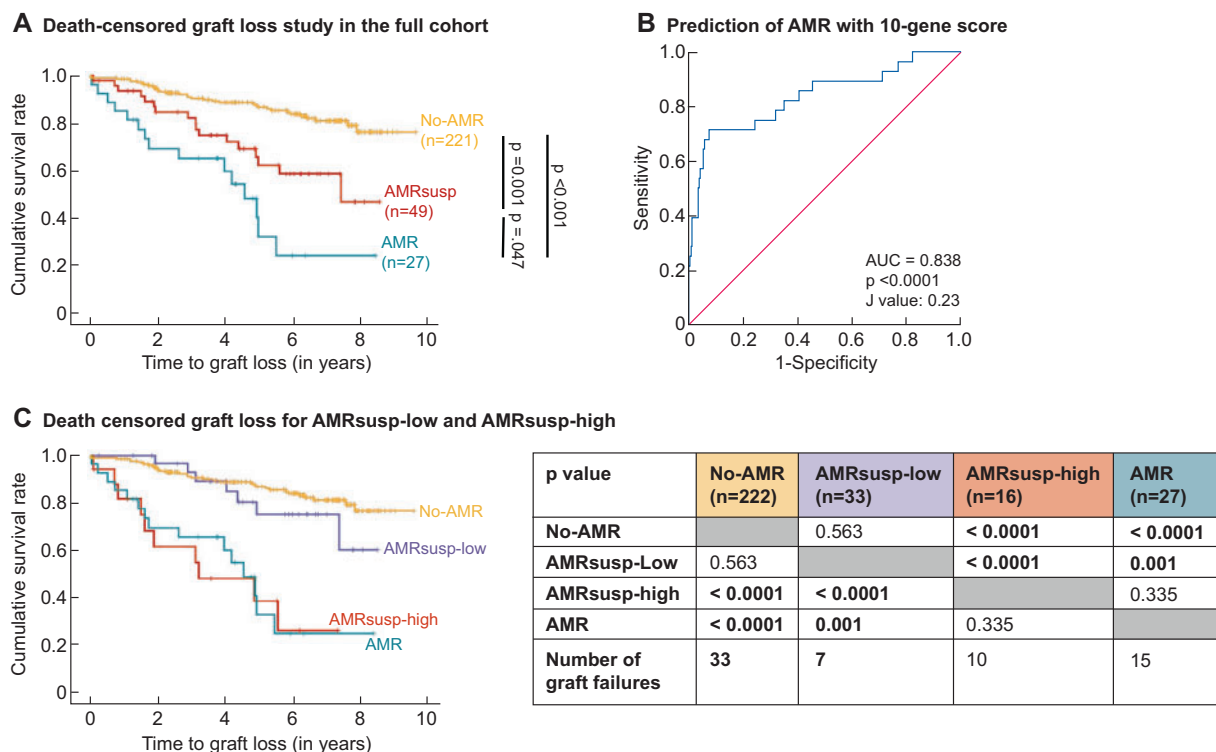
**DISCUSSION**

Gene expression analysis in the diagnosis of AMR was introduced in the 2013 revision of the Banff Classification for Allograft Pathology, but its context of use has never been tested. Most centres still use ‘standard of care’ (SOC) in analysis of transplant biopsies, i.e. histological analysis interpreted in the context of serological data for a DSA. In this study, we compared SOC with SOC + the addition of an AMR 10-gene score, with GL as an outcome. We found improved outcome prediction when adding the AMR 10-gene data to histological and serological data, but not in all biopsies. The 10-gene score only

improved prediction of GL in AMRsusp cases, i.e. cases that show histological evidence of antibody-mediated injury (Criteria 1), but that would only fulfil Criteria 2 (evidence of recent/current interaction of antibody with endothelium) and Criteria 3 (evidence of a circulating DSA) on the basis of gene expression data.

We re-visited the histological and serological characteristics of this group of AMRsusp cases, to see if we could identify particular features in cases with a high AMR 10-gene score compared with those without. Biopsies with a high gene score showed significantly more interstitial fibrosis (ci)/ct (IFTA), a higher ti and a higher ptc score. We suspect this may be related to the fact AMRsusp-high cases showed a trend to being taken at a later time post-transplant, and may indicate that cut-off Banff lesion scores for diagnosis of AMR could be time-sensitive, e.g. for biopsies taken later post-transplant, the selection of a lower ptc score cut-off score for a diagnosis of AMR might increase diagnostic accuracy.

In cases that were clearly AMR using SOC only (i.e. our ‘AMR’ group) and in cases that would never be AMR, even including transcript data (i.e. our ‘No-AMR’ group), the AMR



**FIGURE 2: Outcome analysis.** (A) Outcome analysis was performed using Kaplan–Meier survival curves. Graft survival (in years) from time of biopsy was plotted for No-AMR (orange line), AMRsusp (red line) and AMR (green line). Data were censored for patient death with functioning graft. Statistical analysis was carried out using a Breslow test. (B) ROC curve analysis for the AMR 10-gene score predicting AMR diagnosis, revealed as an AUC of 0.838 (95% CI 0.749–0.943,  $P < 0.0001$ ). Youden’s J statistic was used to define the cut off for the AMR 10-gene score as  $>0.23$ . (C) The cut-off value was used to split the AMRsusp group into AMRsusp-low (AMRsusp samples with an AMR 10-gene score  $<0.23$ ) and AMRsusp-high (AMRsusp samples with an AMR 10-gene score  $>0.23$ ). Outcome analysis was performed using Kaplan–Meier survival curves. Graft survival from time of biopsy (in years) was plotted for No-AMR (orange line), AMRsusp-low (blue line), AMRsusp-high (red line) and AMR (green line). Data were censored for patient death with functioning graft. Statistical analysis was carried out using a Breslow test (results presented in the table).

10-gene score does not appear to improve outcome prediction. This confirms the importance of defining the ‘context of use’ of a biomarker before its validation and clinical implementation [19]. Although applying gene classifiers to unselected cohorts has a central role in biomarker discovery, translation to clinical practice requires a different approach that takes into consideration net clinical benefit and cost-effectiveness compared with SOC. This is particularly important in transplant biopsies, where tissue is scarce, and best use of remaining tissue is dictated by the initial histological findings. For example, elements of the SOC approach that are important to patient management but that cannot be identified using molecular diagnosis include recurrent glomerulonephritis, the degree of scarring and vascular features.

The use of qRT-PCR on a selected gene panel could help disseminate the practice of gene expression analysis, as this is a cheap technology for which most routine diagnostic laboratories have equipment and expertise. It can be performed rapidly and on a single sample at a time if needed. However, it requires a sample taken in RNA preservative to yield large amounts of good quality RNA and its introduction into routine diagnosis would require the introduction of synthetic standards to support inter-laboratory reproducibility.

Our gene selection was based on literature review and confirmed association with AMR in our cohort. The 10-gene panel selected for association with a diagnosis of AMR was then analysed for its association with outcome and found to be predictive of GL. The 10 AMR-associated genes are also present in other AMR-associated gene lists such as the Banff Human Organ Transplant (B-HOT) panel used for Nanostring [20] and the pathogenesis-based transcript sets (core transcript set group DSAT and NKB) for microarray studies [21]. This approach mirrors previous attempts to establish biomarkers for AMR that are both diagnostic (based on their association with serological and histological features) and prognostic (based on their association with outcome). Our panel may not represent the optimal panel for each of these aims. Future multicentre studies on larger unselected cohorts will inform optimal gene selections for defined purposes, including gene weighting. We did not test in this study whether the AMR 10-gene score is applicable using other techniques, or other sample types, such as formalin-fixed paraffin-embedded tissue. However, in another manuscript where we investigated whether results of gene expression analysis performed using qRT-PCR on samples preserved in RNAlater correlated with results using Nanostring nCounter analysis on formalin-fixed paraffin-embedded tissue, these 10 genes correlated well [15]. We previously observed that

**Table 3. Uni- and multivariate analysis of serological, histological and molecular features for predicting GL in the whole cohort**

Variable	Univariate analysis		Multivariate analysis	
	HR (95% CI)	P-value	HR (95% CI)	P-value
<b>General parameters</b>				
Recipient age at time of transplant (continuous)	1.002 (0.985–1.020)	0.815	–	–
Sex (reference = F)	0.717 (0.440–1.169)	0.182	–	–
Type of transplant (reference = LD)	1.602 (0.987–2.600)	0.057	–	–
ABOi (reference = ABO compatible)	0.774 (0.281–2.127)	0.619	–	–
Type of biopsy (reference = surveillance)	<b>2.783 (1.327–5.836)</b>	<b>0.007</b>	<b>2.492 (1.102–5.618)</b>	<b>0.028</b>
Time from transplant to biopsy (continuous)	<b>1.060 (1.016–1.105)</b>	<b>0.007</b>	1.011 (0.960–1.064)	0.679
<b>Banff parameters</b>				
DSA (reference = no DSA, <i>n</i> = 267)				
Class I only present ( <i>n</i> = 11)	0.464 (0.064–3.356)	0.447	–	–
Class II present ( <i>n</i> = 21)	<b>3.014 (1.733–6.723)</b>	<b>&lt;0.001</b>	1.195 (0.329–4.340)	0.787
AMR group (reference = No-AMR, <i>n</i> = 221)				
AMRsusp ( <i>n</i> = 49)	<b>2.833 (1.578–5.084)</b>	<b>&lt;0.001</b>	1.817 (0.746–4.421)	0.188
AMR ( <i>n</i> = 27)	<b>6.012 (3.237–11.164)</b>	<b>&lt;0.001</b>	2.916 (1.102–5.638)	0.085
MI (reference = 0)				
MI = 1 ( <i>n</i> = 42)	<b>2.358 (1.212–4.587)</b>	<b>0.011</b>	1.307 (0.525–3.254)	0.564
MI > 1 ( <i>n</i> = 40)	<b>5.626 (3.243–9.762)</b>	<b>&lt;0.001</b>	<b>2.543 (1.035–6.252)</b>	<b>0.042</b>
C4d (any positive <i>n</i> = 40; reference = C4d0)				
t (any positive <i>n</i> = 87; reference = t0)	1.137 (0.697–1.852)	0.608	–	–
i (any positive <i>n</i> = 46; reference = i0)	1.179 (0.617–2.254)	0.619	–	–
v (any positive <i>n</i> = 5; reference = v0)	<b>3.923 (1.229–12.519)</b>	<b>0.021</b>	1.309 (0.363–4.714)	0.681
IFTA% (continuous)	<b>1.030 (1.020–1.041)</b>	<b>&lt;0.001</b>	<b>1.029 (1.015–1.043)</b>	<b>&lt;0.001</b>
cg (any positive <i>n</i> = 31; reference = cg0)	<b>4.201 (2.277–7.750)</b>	<b>&lt;0.001</b>	1.044 (0.433–2.517)	0.924
Cut off 10-gene AMR score (reference ≤ 0.23)	<b>2.389 (1.462–3.903)</b>	<b>0.001</b>	1.094 (0.446–2.683)	0.844
10-gene AMR score (continuous)	<b>1.091 (1.057–1.126)</b>	<b>&lt;0.001</b>	1.026 (0.942–1.119)	0.554

Cox proportional hazards regression analysis on 297 samples. Abbreviations: HR, Hazard Ratio, CI, Confidence Interval; LD, Live Donor, ABO, blood group; AMR antibody-mediated rejection; C4d, Complement component 4; IFTA, Interstitial fibrosis/tubular atrophy; t, tubulitis; i, interstitial inflammation; v, intimal arteritis; MI, microcirculation; cg, glomerular capillary wall double contours. Bold values, denote statistical significance at the *P* < 0.05.

**Table 4. Uni- and multivariate analysis of serological, histological and molecular features for predicting GL in the AMRsusp (*n* = 49)**

Variable	Univariate analysis		Multivariate analysis	
	HR (95% CI)	P-value	HR (95% CI)	P-value
<b>General parameters</b>				
Recipient age at time of transplant (continuous)	0.989 (0.963–1.015)	0.408	–	–
Sex (reference = F)	0.807 (0.309–2.103)	0.660	–	–
Type of transplant (reference = LD)	2.510 (0.919–6.855)	0.073	–	–
ABOi (reference = ABO compatible)	0.876 (0.115–6.656)	0.899	–	–
Type of biopsy (reference = surveillance)	2.351 (0.660–8.372)	0.187	–	–
Time from transplant to biopsy (continuous)	<b>1.069 (1.008–1.134)</b>	<b>0.025</b>	1.012 (0.941–1.088)	0.752
<b>Banff parameters</b>				
DSA (reference = no DSA)				
NA				
MI (reference = 0)				
MI = 1 ( <i>n</i> = 23)	0.720 (0.193–2.685)	0.625	–	–
MI > 1 ( <i>n</i> = 14)	2.257 (0.656–7.761)	0.196	–	–
C4d (any positive <i>n</i> = 15; reference = C4d0)				
t (any positive <i>n</i> = 14; reference = t0)	1.032 (0.362–2.941)	0.953	–	–
i (any positive <i>n</i> = 8; reference = i0)	1.834 (0.596–5.645)	0.290	–	–
v (any positive <i>n</i> = 2; reference = v0)	1.394 (0.175–10.980)	0.758	–	–
IFTA% (continuous)	<b>1.034 (1.013–1.055)</b>	<b>0.001</b>	<b>1.032 (1.005–1.060)</b>	<b>0.022</b>
cg (any positive <i>n</i> = 12; reference = cg0)	1.587 (0.558–4.517)	0.387	–	–
Cut off 10-gene AMR score (reference ≤ 0.23)	<b>3.844 (1.426–10.363)</b>	<b>0.008</b>	1.292 (0.216–7.734)	0.779
10-gene AMR score (continuous)	<b>1.109 (1.033–1.191)</b>	<b>0.004</b>	<b>1.138 (1.029–1.286)</b>	<b>0.012</b>

Cox proportional hazards regression analysis on 49 AMRsusp samples. Abbreviations: HR, Hazard Ratio, CI, Confidence Interval; LD, Live Donor, ABO, blood group; AMR antibody-mediated rejection; C4d, Complement component 4; IFTA, Interstitial fibrosis/tubular atrophy; t, tubulitis; i, interstitial inflammation; v, intimal arteritis; MI, microcirculation; cg, glomerular capillary wall double contours; NA, not applicable as only two DSA-positive samples. Bold values, denote statistical significance at the *P* < 0.05.

there was good correlation between measurements of AMR-associated genes, when comparing two different cores from the same kidney transplant, using two different techniques (Nanostring and qRT-PCR) [15].

This study has limitations. First, we used a retrospective cohort limited to 297 samples, selected based on the availability of sufficient quantities of cDNA for qRT-PCR analysis. Although it contains a mix of diagnoses that are representative of our



biopsy population, we cannot exclude a sample bias. Second, this is a retrospective cohort with varied therapeutic strategies, so we were not able to capture response to AMR treatment as an outcome. Prospective cohort analyses, preferably in the context of clinical trials of AMR, are needed. We used the commonly accepted graft failure as our outcome measure. GL has many causes other than AMR; the study design and balanced demographic and transplant data between the groups only go part of the way to correcting for this limitation [22].

In summary, we provide evidence for the first time that an AMR 10-gene score can improve prediction of graft outcome in the context of use proposed in the Banff Classification for Allograft Pathology. Larger, multicentre prospective studies will be needed to better define optimal gene scores for a variety of gene expression analysis techniques.

## SUPPLEMENTARY DATA

Supplementary data are available at [ndt online](https://www.ndt-online.org).

## ACKNOWLEDGEMENTS

We thank Dr Ben Adam for sharing R scripts for logistic regression analysis, and for helpful comments. Human samples used in this research project were obtained from the Imperial College Healthcare Tissue Bank (ICHTB). ICHTB is supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London.

## FUNDING

F.T. and C.R. are supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London. C.R.'s research activity is made possible with generous support from Sidharth and Indira Burman. Infrastructure support for this research was provided by the NIHR Imperial Biomedical Research Centre.

## AUTHORS' CONTRIBUTIONS

C.R., A.M. and F.T. conceived and designed the project. C.R., A.M., H.T.C., F.T., J.B. and M.W. recorded and retrieved clinical and pathological data. E.S. recorded and retrieved tissue typing data. F.T., K.D. and C.R. performed the sample analysis and analysed the data. All authors contributed to writing the manuscript.

## CONFLICT OF INTEREST STATEMENT

C.R. has a consultancy agreement with UCB.

## REFERENCES

- Halloran PF, Famulski KS, Reeve J. Molecular assessment of disease states in kidney transplant biopsy samples. *Nat Rev Nephrol* 2016; 12: 534–548

- Sis B, Jhangri GS, Bunnag S *et al*. Endothelial gene expression in kidney transplants with alloantibody indicates antibody-mediated damage despite lack of C4d staining. *Am J Transplant* 2009; 9: 2312–2323
- Hidalgo LG, Sis B, Sellares J *et al*. NK cell transcripts and NK cells in kidney biopsies from patients with donor-specific antibodies: evidence for NK cell involvement in antibody-mediated rejection. *Am J Transplant* 2010; 10: 1812–1822
- Yazdani S, Callemeyn J, Gazut S *et al*. Natural killer cell infiltration is discriminative for antibody-mediated rejection and predicts outcome after kidney transplantation. *Kidney Int* 2019; 95: 188–198
- Haas M, Loupy A, Lefaucheur C *et al*. The Banff 2017 kidney meeting report: revised diagnostic criteria for chronic active T cell-mediated rejection, antibody-mediated rejection, and prospects for integrative endpoints for next-generation clinical trials. *Am J Transplant* 2018; 18: 293–307
- Hayde N, Bao Y, Pullman J *et al*. The clinical and molecular significance of C4d staining patterns in renal allografts. *Transplantation* 2013; 95: 580–588
- Hayde N, Broin PO, Bao Y *et al*. Increased intragraft rejection-associated gene transcripts in patients with donor-specific antibodies and normal biopsies. *Kidney Int* 2014; 86: 600–609
- Dominy KM, Roufosse C, de Kort H *et al*. Use of quantitative real time polymerase chain reaction to assess gene transcripts associated with antibody-mediated rejection of kidney transplants. *Transplantation* 2015; 99: 1981–1988
- Dominy KM, Willicombe M, Al Johani T *et al*. Molecular assessment of C4d-positive renal transplant biopsies without evidence of rejection. *Kidney Int Rep* 2019; 4: 148–158
- Adam B, Afzali B, Dominy KM *et al*. Multiplexed color-coded probe-based gene expression assessment for clinical molecular diagnostics in formalin-fixed paraffin-embedded human renal allograft tissue. *Clin Transplant* 2016; 30: 295–305
- Sellares J, Reeve J, Loupy A *et al*. Molecular diagnosis of antibody-mediated rejection in human kidney transplants. *Am J Transplant* 2013; 13: 971–983
- Loupy A, Lefaucheur C, Vernerey D *et al*. Molecular microscope strategy to improve risk stratification in early antibody-mediated kidney allograft rejection. *J Am Soc Nephrol* 2014; 25: 2267–2277
- Roufosse C, Simmonds N, Clahsen-van Groningen M *et al*. A 2018 reference guide to the Banff classification of renal allograft pathology. *Transplantation* 2018; 102: 1795–1814
- Loupy A, Haas M, Solez K *et al*. The Banff 2015 kidney meeting report: current challenges in rejection classification and prospects for adopting molecular pathology. *Am J Transplant* 2017; 17: 28–41
- Toulza F, Dominy K, Cook T *et al*. Technical considerations when designing a gene expression panel for renal transplant diagnosis. *Sci Rep* 2020; 10: 17909
- Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(Delta Delta C(T)) method. *Methods* 2001; 25: 402–408
- de Hoon MJ, Imoto S, Nolan J *et al*. Open source clustering software. *Bioinformatics* 2004; 20: 1453–1454
- Roufosse C, Drachenberg C, Renaudin K *et al*. Molecular assessment of antibody-mediated rejection in human pancreas allograft biopsies. *Clin Transplant* 2020; 34: e14065
- Naesens M, Friedewald J, Mas V *et al*. A practical guide to the clinical implementation of biomarkers for subclinical rejection following kidney transplantation. *Transplantation* 2020; 104: 700–707
- Mengel M, Loupy A, Haas M *et al*. Banff 2019 meeting report: molecular diagnostics in solid organ transplantation-consensus for the Banff Human Organ Transplant (B-HOT) gene panel and open source multicenter validation. *Am J Transplant* 2020; 20: 2305–2317
- Halloran PF, de Freitas DG, Einecke G *et al*. The molecular phenotype of kidney transplants. *Am J Transplant* 2010; 10: 2215–2222
- Einecke G, Reeve J, Gupta *et al*. INTERCOMEX Investigators. Factors associated with kidney graft survival in pure antibody-mediated rejection at the time of indication biopsy: importance of parenchymal injury but not disease activity. *Am J Transplant* 2021; 21: 1391–1401

Received: 12.3.2021; Editorial decision: 17.6.2021