Transplant International

ORIGINAL ARTICLE

Microvascular inflammation in renal allograft biopsies assessed by endothelial and leukocyte co-immunostain: a retrospective study on reproducibility and clinical/prognostic correlates

Marco Delsante[1,*], Umberto Maggiore[2,*], Jonathan Levi[3], David E. Kleiner[3], Annette M. Jackson[4], Lois J. Arend[1], Stephen M. Hewitt[3], Naima Carter-Monroe[5], Serena M. Bagnasco[1] & Avi Z. Rosenberg[1]

1 Renal Pathology Service, Johns Hopkins University, Baltimore, MD, USA
2 Kidney-Pancreas Transplant Unit, Department of Nephrology, Parma University Hospital, Parma, Italy
3 Laboratory of Pathology, National Cancer Institute, Bethesda, MD, USA
4 Immunogenetics Laboratory, Johns Hopkins University, Baltimore, MD, USA
5 VA Medical Center, Baltimore, MD, USA

Correspondence
Marco Delsante, Department of Nephrology, Parma University Hospital, Via Gramsci 15, 43125 Parma, Italy.
Tel.: +39 3293166082;
fax: +39 0521702810;
e-mail: mdelsan1@jhmi.edu
Avi Z. Rosenberg, Pathology, Johns Hopkins University, 600 North Wolfe Street, Baltimore, MD 21231, USA.
Tel.: +1 9179419845;
fax: +1 410-614-7110;
e-mail: arosen34@jhmi.edu

*These authors contributed equally.

SUMMARY

The most prominent histologic lesion in antibody-mediated rejection is microvascular inflammation (MVI); however, its recognition and scoring can be challenging and poorly reproducible between pathologists. We developed a dual immunohistochemical (IHC)-stain (anti-CD34/anti-CD45 for endothelium/leukocytes) as ancillary tool to improve on the semi-quantitative Banff scores and allow quantification of MVI. We examined the relationship between CD34–CD45 IHC-based quantitative MVI score (the inflamed peritubular capillary ratio, iptcr) and renal-graft failure or donor-specific antibodies (DSA) strength at the time of biopsy. Quantitative iptcr score was significantly associated with renal graft failure (hazard ratio 1.81, per 1 SD-unit [0.13 points] of iptcr-increase; $P = 0.026$) and predicted the presence and strength of DSA (ordinal odds ratio: 2.42; $P = 0.005$; 75 biopsies/60 kidney transplant recipients; 30 HLA- and/or ABO-incompatible). Next, we assessed inter-pathologist agreement for ptc score and ptc extent (focal/diffuse) using CD34–CD45 IHC as compared to conventional stain. Compared to conventional stain, CD34–CD45 IHC significantly increased inter-pathologist agreement on ptc score severity and extent (κ-coefficient from 0.52–0.80 and 0.46–0.68, respectively, $P < 0.001$). Our findings show that CD34–CD45 IHC improves reproducibility of MVI scoring and facilitates MVI quantification and introduction of a dual anti-CD34/CD45 has the potential to improve recognition of MVI ahead of DSA results.

Transplant International 2019; 32: 300–312

Key words
kidney clinical, rejection

Received: 18 June 2018; Revision requested: 13 July 2018; Accepted: 31 October 2018; Published online: 11 December 2018

Introduction

Antibody mediated rejection (AMR) is a major cause of long-term renal allograft loss [1] and results from the development of donor-specific antibodies in the transplant recipient against class I or class II human leukocyte antigens (HLA-DSA) or against non-HLA proteins [2]. Donor-specific antibodies (DSA) activate the complement cascade by binding to antigens on allograft endothelial cells, resulting in endothelial C4d

doi:10.1111/tri.13371

deposition and microvascular inflammation (MVI). MVI is defined as the recruitment of leukocytes to glomerular (glomerulitis – g) and/or peritubular capillaries (peritubular capillaritis – ptc) [3]. The assessment of these features in the biopsy is critical for the diagnosis of AMR. Poor reproducibility of MVI scores [4–9], limitations of C4d sensitivity and heterogeneity in DSA testing standards have spurred an interest in precise assessment of microvascular inflammation [10,11]. That represents a serious concern in clinical practice, when rapid clinical decision solely based on the biopsy findings may be required, DSA results not being readily available. Here we present a combined immunohistochemical stain for CD34 (vascular endothelium) and CD45 (Leukocyte Common Antigen – LCA) that allows for the quantitative assessment of microvascular inflammation enhancing the spatial resolution of the pattern of inflammation. Our study tested two hypotheses: (i) CD34–CD45 immunohistochemistry (IHC) stain-derived quantitative microvascular inflammation scores correlates with (a) serologic DSA testing results and (b) predicts clinical endpoints (i.e., graft failure) compared to semi-quantitative Banff scores, and (ii) CD34–CD45 IHC stain assessment of classic Banff ptc score can improve inter-pathologist agreement compared to assessment by routine histologic stains.

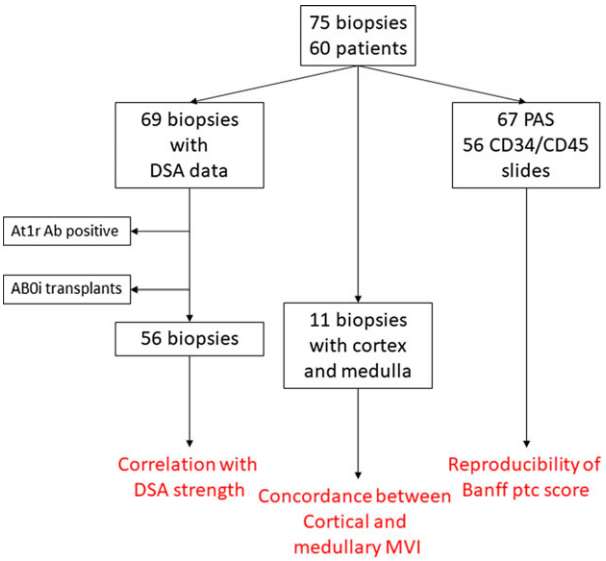

Figure 1 Study data flow chart. Of the initial pool of 75 biopsies, 69 had available DSA data; after excluding At1r-antibodies positive patients and ABOi transplants, 56 biopsies were used to assess correlation between quantitative MVI scores and DSA strength. Eleven biopsies had both cortex and medulla and allowed for discrete regional evaluation of quantitative ptc inflammation. Finally, for reproducibility studies, only glass slides where allowed for inter-pathologist circulation. From the initial group of 75 biopsies, 67 had available PAS and 56 had CD34/CD45 glass slides which were then circulated between the four renal pathologists. ABOi, ABO-incompatible transplant; DSA, donor-specific anti-HLA antibodies; MVI, microvascular inflammation.

Materials and methods

Study population and biopsies

We retrospectively evaluated a cohort of 75 protocol or for-cause renal allograft biopsies (from 60 patients) performed at The Johns Hopkins Hospital between 2004 and 2016. Allograft biopsies were selected for CD34–CD45 IHC staining to include the most common conditions which are associated with variable degrees of MVI and, therefore, that must be distinguished from antibody mediated rejection (Fig. 1; Appendix S1, Table S1). Allograft biopsies were scored according to the 2013 Banff Classification [12]. We retrospectively extracted clinical data for the biopsied patients. The study was approved by the Johns Hopkins Medicine Institutional Review Board (IRB number 00090103).

Immunohistochemistry staining and microscopy

Immunohistochemistry stain was performed using commercially available, clinical grade antibodies and detection kit. IHC for endothelial cells was performed using monoclonal mouse-anti-human CD34 antibody (Ventana, Tucson, AZ, USA, 790-2927) and for leukocytes using monoclonal mouse-anti-human anti-CD45 (Dako, Carpintera, CA, USA, M0701) on 4 μm formalin-fixed paraffin-embedded tissue sections using a Ventana Benchmark Ultra auto-stainer (Roche, Basel, Switzerland). CD34 was chosen based on studies showing a more diffuse positivity in human kidney vessels compared to other endothelial markers such CD31 [13]. Detection was performed using UltraView alkaline phosphatase red detection for CD34 (Ventana, 760-501, Red) and UltraView DAB detection for CD45 (Ventana, 760-500, Brown). Examples of the observed staining with various degrees of MVI are shown in Figs 2 and 3. Stained slides were digitally scanned (NanoZoomer 2.0HT, Hamamatsu, Japan) and analyzed using NDPVIEW2 software (www.hamamatsu.com). A PAS stained section was evaluated as the comparator for the current Banff scoring system of MVI.

Quantitative microvascular inflammation scores

We manually assessed quantitative microvascular inflammation in the different compartment of renal

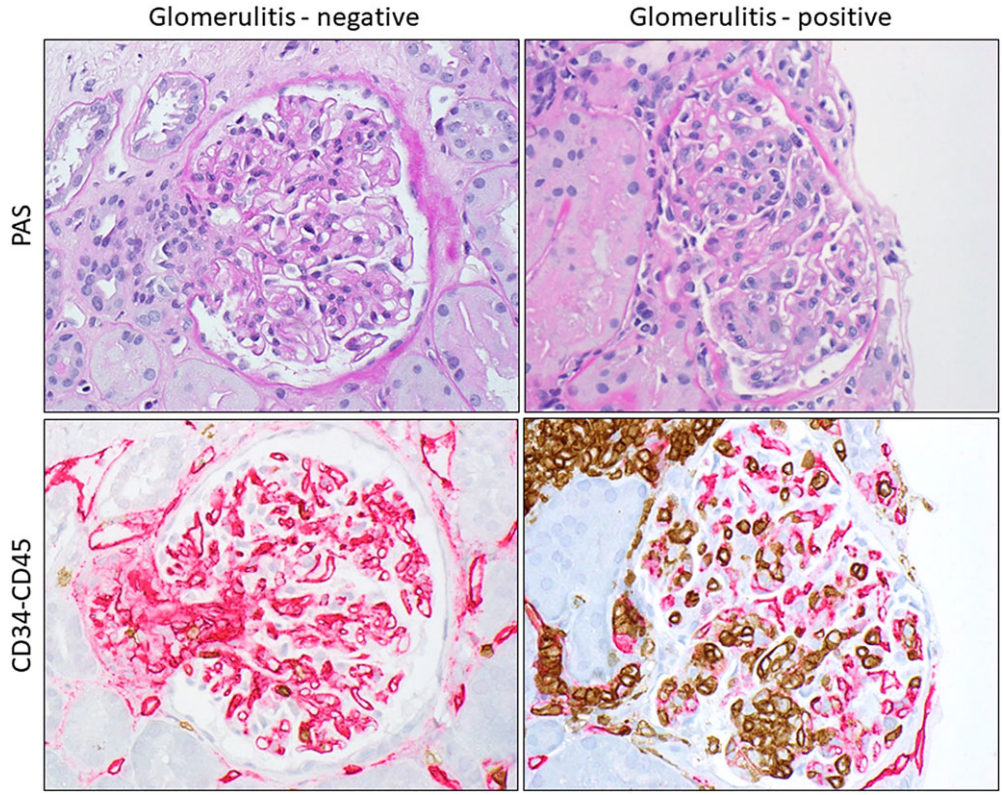

Figure 2 Glomerular inflammatory cells infiltrate (glomerulitis, Banff g score) with PAS (upper panels) and CD34–CD45 immunohistochemical (bottom panels) stain. Magnification 400×.

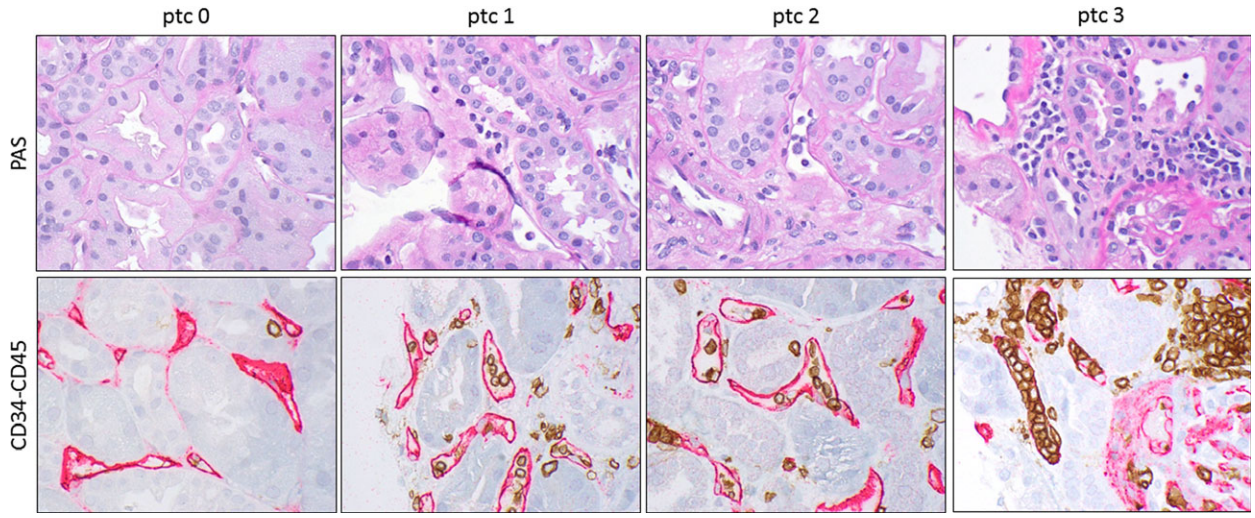

Figure 3 Peritubular capillary inflammatory cells margination (peritubular capillaritis, Banff ptc score) with PAS (upper panels) and CD34–CD45 immunohistochemical (bottom panels) stain. Magnification 600×.

biopsies for each scanned CD34–CD45 IHC using the counting tool in NDPview2 (Table 1). Glomerular quantitative scores included: (i) mean leukocytes per glomerulus (gml – ratio between number of glomerular leukocytes and total number of glomeruli) and (ii) number of leukocytes observed in the most inflamed glomerulus (gmaxl). For peritubular capillary inflammation the following parameters were evaluated: (i) mean leukocytes per peritubular capillary (ptcml – the ratio between total number of leukocytes in peritubular capillaries and total number of peritubular capillaries), (ii) number of leukocytes observed in the most inflamed

Table 1. Microvascular inflammation scores.

Score	Abbreviation	Definition
Banff scores		
Glomerulitis	g	Glomerular mononuclear cell infiltrate and endothelial cell enlargement g1: <25% of glomeruli g2: 25–75% of glomeruli g3: >75% of glomeruli
Peritubular capillaritis	ptc	Mononuclear cell margination in peritubular capillaries ptc0: <10% ptc involved ptc1: max 3–4 cells/ptc in ≥10% ptc ptc2: max 5–10 cells/ptc in ≥10% ptc ptc3: max >10 cells/ptc in ≥10% ptc
Peritubular capillaritis extension	ptc extension	Focal ptc: <50% of cortical peritubular capillaries Diffuse ptc: >505 of cortical peritubular capillaries
Quantitative MVI scores		
Glomerular mean leukocytes	gml	$\frac{\text{number of glomerular leukocytes}}{\text{number of glomeruli}}$
Glomerular max leukocytes	gmaxl	Number of leukocytes observed in the most inflamed glomerulus per biopsy
Peritubular capillary mean leukocytes	ptcml	$\frac{\text{number of peritubular leukocytes}}{\text{number of peritubular capillaries}}$
Peritubular capillary max leukocytes	ptcmaxl	Number of leukocytes observed in the most inflamed peritubular capillary
Inflamed peritubular capillaries ratio	iptcr	$\frac{\text{number of peritubular capillaries with} \geq 1 \text{ leukocyte}}{\text{total number of peritubular capillaries}}$

peritubular capillary (ptcmaxl) and (iii) inflamed peritubular capillaries ratio (iptcr – the ratio between number of peritubular capillaries with at least one marginated leukocyte and the total number of peritubular capillaries). All available biopsies showing cortical and medullary tissue (11 samples) where assessed for quantitative ptc scores in the different renal tissue compartments.

Reproducibility studies

Sixty-seven PAS stained biopsies and 56 CD34–CD45 IHC stained biopsies were independently reviewed by four renal pathologists (MD, AZR, SMB and DEK) to assess the inter-pathologist agreement of Banff ptc scores following the Banff schema [14]. Pathologists were blinded to clinical information. PAS and CD34–CD45 IHC stain were reviewed separately.

Antibody tests

HLA-specific antibodies were evaluated as part of the standard clinical care approach using solid-phase immunoassays (Lifecodes classes I and II phenotype panels; Immucor-Lifecodes, Stamford, CT, USA and Single Antigen Beads; One Lambda, Canoga Park, CA, USA) performed on a Luminex platform. The majority of sera (>90%) were treated to remove interfering substances [15] and tests of serum dilutions were performed as needed. HLA-DSA was correlated with crossmatch strength as previously described [16] using cumulative Luminex bead data and categorized by three thresholds: (i) CDC-XM level, (ii) Flow-XM level and (iii) Luminex positive level. Briefly, MFI ranges for HLA-A, -B and –DR were ≥10 000 for CDC-XM level, 4000–9000 for Flow-XM level and 1000–4000 for Luminex level. Such categories have been found to hold clinical relevance in the Johns Hopkins positive crossmatch cohort and when examining HLA incompatible transplant outcomes across US centers [17,18]. Additionally, for HLA-C, -DQ, and –DP, CDC-XM and Flow-XM levels were determined using 1:32 and 1:8 serum dilutions, respectively. MFI values below 1000 and lacking HLA specificity patterns were reported as negative. Table 3 contains the single antigen panel MFI data for the summation of all detected HLA-DSA at time of biopsy. Detection of antibodies against angiotensin II type 1 receptor (AT1R-Ab) was performed using quantitative ELISA (CellTrend GmbH,

Luckenwalde, Germany) as previously described using a positive threshold of 17 U/ml [19]. Sera collected at the time of the biopsy from five patients were tested for AT1R antibody in response to antibody-mediated injury observed on biopsy in the absence of HLA-DSA.

Statistical analysis

We measured agreement between the four pathologists for peritubular capillaritis and extent of peritubular capillary inflammation assessment, by calculating the kappa

Table 2. Characteristics of the study population

Variables	Results	Results	Results
Number of patients	60		
Gender			
Male	31 (51.6)		
Female	29 (48.4)		
Race			
Caucasian	33 (55)		
African American	19 (31.6)		
Asian	2 (3.3)		
Hispanic	1 (1.6)		
Other	5 (8.3)		
Primary kidney disease			
Diabetes	10 (16.7)		
Hypertension	11 (18.4)		
Glomerulonephritis	16 (26.7)		
ADPKD	4 (6.7)		
CAKUT	7 (11.6)		
Other	6 (10.0)		
Unknown	6 (10.0)		
Age at transplant, years	45 ± 17		
Donor type			
Deceased	24 (40)		
Living related	17 (28.4)		
Living unrelated	19 (31.6)		
Immunologic risk*			
Repeated transplants	19 (31.6)	2 transplants	16 (26.7)
		3 transplants	2 (3.3)
		4 transplants	1 (1.6)
HLA-incompatible	25 (36.6)		
ABO-incompatible	2 (3.3)		
HLA- and ABO-incomp.	3 (5.0)		
Immunosuppressive treatment			
Desensitization	27 (45)	PEX	8 (29.6)
		PEX + RTX	12 (44.4)
		PEX + IVIG	2 (7.4)
		PEX + RTX + IVIG	2 (7.4)
		RTX	3 (11.1)
Induction	48 (80)	DACLIZUMAB	9 (18.7)
		THYMOGLOBULIN	28 (58.3)
		BASILIXIMAB	7 (14.5)
		THYMO + ECULIZUMAB	2 (4.1)
		RTX	1 (2.0)
		ALEMTUZUMAB	1 (2.0)
Maintenance	60 (100)	PDN, TAC, MMF	54 (90)
		PDN, TAC	1 (1.6)
		TAC	1 (1.6)
		PDN, TAC, MMF, ECULIZUMAB	1 (1.6)
		PDN, CYC, MMF	2 (3.3)
		PDN, TAC, SIR	1 (1.6)

Transplant International 2019; 32: 300–312

Table 2. Continued.

Variables	Results	Results	Results
Follow-up, months	18 (6–371)		
eGFR, ml/min/1.73 m^2			
Baseline	56.5 ± 27.8		
Follow-up	38.4 ± 29.6		
Graft loss	12 (20)		

Categorical variables are reported as number (percentage), continuous variables as mean ± standard deviation or median (range).

ADPKD, autosomal dominant polycystic kidney disease; CAKUT, congenital anomalies of kidney and urinary tract; CYC, cyclophosphamide; IVIG, low-dose intravenous immunoglobulins; MMF, mycofenolate mofetil; PDN, prednisone; PEX, plasma exchange; RTX, rituximab; SIR, sirolimus; TAC, tacrolimus.

*Some of the patients who received repeated kidney transplantations were also ABO-incompatibile or ABO- and HLA-incompatible.

agreement coefficient with Gwet's method (see Appendix S1) [20]. For each agreement coefficient we computed the 95% confidence interval and the paired statistical test of the difference between the agreement coefficients using conventional as compared to CD34–CD45 IHC stain [21]. To assess the correlation between cortical and medullary iptcr, we used the Lin's concordance coefficient [16]. We examined the relationship between graft failure (defined as progression to chronic renal replacement therapy or reduction in eGFR below 10 ml/min/1.73 m^2) and each histologic score by fitting for each histological score a separate Cox proportional-hazards regression model adjusted for antibody strength and with the time at risk starting on the date of the first biopsy, and the histologic score being included as time-dependent variable. The relationship between DSA strength category and each histologic score was fitted using ordinal logistic regression models. Biopsies from patients with negative HLA-DSA test but positive AT1R antibodies as well as from recipients of ABO-incompatible transplant were not included in this analysis. Reported confidence intervals and *P*-values were calculated using robust standard errors to account for repeated biopsies in each patient [22]. Because of the presence of repeated biopsies, the analyses were carried out using the appropriate statistical tools accounting for the correlated nature of the data. To visualize the ordinal regression model, we plotted the fitted predicted probability for HLA-DSA presence and strength against the measured inflamed peritubular capillaries ratio (iptcr). To simplify the comparison between the histologic score regression coefficients, all the histologic score variables were standardized and for every regression model, coefficients represent the change in the outcome variable per one standard deviation unit increase in the histologic score. STATA SE release 15.0 (2017, Stata Corp, College Station, TX, USA) was used for all the analyses. A more detailed description of the statistical analysis is provided in the Appendix S1.

Results

Study population

The cohort included 60 patients (31 males, 29 females) for a total of 75 biopsies. Patients' characteristics are reported in Table 2. The most frequent cause of end stage renal disease (ESRD) was primary or secondary glomerulonephritis (26.6% of cases), followed by hypertensive glomerulosclerosis and diabetic nephropathy (18.4% and 16.7%, respectively). Twenty-five transplants were HLA-incompatible with HLA-DSA at the time of the transplant; two were ABO-incompatible and three were both HLA- and ABO-incompatible. None of the ABO-incompatible recipients had blood group antibody titers >8 at the time of the biopsy. The median follow-up was 31 months (range 1–362) with 20% graft loss. Mean eGFR was 56.5 ± 27.8 ml/min/1.73 m^2 at baseline and 38.4 ± 29.6 ml/min/1.73 m^2 at last available follow-up. No evidence of rejection was present in 24% of the cases; isolated AMR alone was diagnosed in 12% of the biopsies (six cases C4d-positive and three cases C4d-negative); 12% had isolated CMR (nine cases: five Banff IA, three IB and one IIA). Findings consistent with borderline for CMR and "suspicious for AMR" were observed in 4% and 18.6% of the cases, respectively. Overall, features of AMR (including AMR, mixed rejections and "suspicious for AMR" cases) were observed in 60.0% of the biopsy, with a median time from transplant of 17 months (range 1–91). The DSA status showed that most of the patients (68.2%) had donor-specific antibodies at the time of the biopsy, directed against HLA in most of the cases, with variable

Table 3. Biopsy characteristics

Variables	Results	Results
Number of biopsies		75*
Mean patent glomeruli per biopsy		13 ± 8.8
Time after transplant, months		31 (1–362)
Age at biopsy, years		48 ± 16
Type of biopsy	Protocol	13 (17.3)
	For cause	57 (76.0)
	Follow-up	5 (6.6)
DSA status at biopsy	HLA-DSA strength	
	Negative	26 (37.6)
	Luminex	19 (27.5)
	Flow-XM	14 (20.2)
	CDC-XM	10 (14.4)
	DSA classes	
	Negative	22 (31.8)
	Class I HLA	11 (15.9)
	Class II HLA	18 (26.0)
	Class I + II HLA	14 (20.2)
	At1r	4 (5.7)
HLA-DSA Sum MFI values at biopsy	Class I HLA-DSA	
	Luminex	2048 (1000–8706)
	Flow-XM	13 824 (12 694–20 409)
	CDC-XM	None
	Class I HLA-DSA	
	Luminex	2671 (1123–9825)
	Flow-XM	17 551 (12 446–26 885)
	CDC-XM	28 487 (22 054–31 277)
eGFR at biopsy, ml/min/1.73 m^2		41.7 ± 25.85
Histological lesions of rejection	Negative	18 (24.0)
	Borderline	3 (4.0)
	CMR	
	IA	5 (6.6)
	IB	3 (4.0)
	IIA	1 (1.3)
	Suspicious for AMR	14 (18.6)
	AMR	
	C4d positive	6 (8.0)
	C4d negative	3 (4.0)
	AMR + borderline	14 (18.6)
	Mixed Rejection	8 (10.6)
Other findings	CNI toxicity	8 (10.6)
	ATI	13 (17.3)
	Recurrent GN	8 (10.6)
	BK-virus nephropathy	2 (2.6)
	Other	10 (13.3)
	None	34 (45.3)

AMR, antibody mediated rejection; ATI, acute tubular injury; AT1R, anti-angiotensin type 1 receptor antibodies; CDC-XM, cytotoxic crossmatch; CMR, T-cell mediated rejection; CNI, calcineurin inhibitor; DSA, donor-specific antibodies; Flow-XM, flow cytometric crossmatch; GN, glomerulonephritis; HLA-DSA, donor-specific HLA antibodies.

Categorical variables are reported as number (percentage), continuous variables as mean ± standard deviation or median (range).

*Due to missing scores, regression analyses were based on a median on 68 biopsies, whereas the comparison between interpathologist agreement coefficients between conventional and IHC CD34–CD45 stain was based on 52 biopsies.

Transplant International 2019; 32: 300–312

Table 4. Association between histologic scores and graft loss.

Scores	Beta coefficient	HR	95% CI	*P* value
Banff scores				
g	−0.22	0.80	0.42–1.54	0.51
ptc	−0.12	0.89	0.46–1.71	0.73
i	0.22	1.24	0.72–2.14	0.43
t	0.43	1.53	0.87–2.69	0.14
v	−0.09	0.91	0.37–2.27	0.84
ah	−0.05	0.96	0.56–1.64	0.87
cg	0.15	1.16	0.73–1.86	0.53
ci	**0.68**	**1.97**	**0.99–3.92**	**0.054**
ct	**0.70**	**2.02**	**1.03–3.96**	**0.040**
cv	−0.07	0.93	0.51–1.71	0.83
mm	0.07	1.07	0.67–1.73	0.77
ti	0.29	1.34	0.77–2.33	0.29
C4d	−0.40	0.67	0.32–1.38	0.27
Quantitative MVI scores				
gml	−0.18	0.83	0.40–0.74	0.63
gmaxl	0.08	1.08	0.62–1.88	0.78
Cortical ptcml	0.32	1.38	0.82–2.32	0.23
Cortical ptcmaxl	0.49	1.64	0.88–3.03	0.12
Cortical iptcr	**0.74**	**2.09**	**1.09–4.01**	**0.026**

95% CI, 95 percent confidence interval; ah, arteriolar hyalinosis; Beta coefficient, log(HR); cg, chronic glomerulopathy; ci, interstitial fibrosis; ct, tubular atrophy; cv, chronic arteriopathy; g, glomerulitis; gmaxl, max glomerular leukocyte; gml, mean glomerular leukocytes; HR, hazard ratio; i, interstitial inflammation; iptcr, inflamed peritubular capillaries ratio; mm, mesangial expansion; MVI, microvascular inflammation; ptcmaxl, max peritubular capillary leukocytes; ptcml, mean peritubular capillary leukocytes; ptc, peritubular capillaritis; ti, total inflammation; t, tubulitis; v, endoarteritis. Bold indicates statistically significant.

Results of separate Cox proportional-hazards regression models adjusted for antibody-strength and examining the relationship between the time-varying variable histologic score (i.e. time-varying multiple biopsies over the follow-up) and graft failure. Hazard ratios represent the relative increase in the hazard of renal graft failure based on the scores of the last biopsy per one standard deviation unit increase in the histologic score. Cortical iptcr represents the increase of the risk per one standard deviation unit (i.e. 0.13) above the threshold value of iptcr = 0.10. Beta coefficients (log HR) are reported along with HRs in order to make easier the direct comparison between each histologic score: the larger the absolute value of the beta coefficient the higher is the risk of ESRD. In fact, because of missing data, number of measurements differed between histologic scores (median 67 [range 51–68]).

antibody strength (Table 3). Examples of glomerulitis and peritubular capillaritis with PAS and CD34–CD45 IHC are shown in Figs 2 and 3.

Correlation between MVI scores and clinical outcome

Quantitative MVI scores were correlated with clinical data including risk of graft failure and characteristics of DSA and the performance was compared with standard Banff scores. After a median follow-up of 18 months from transplant (range 0.4–91) 12 patients experienced graft failure and one patient died with a functioning graft. The results of Cox proportional-hazards regression analysis with time-dependent covariates examining the relationship between the rate of ESRD and histologic scores are reported in Table 4. Hazard ratios are interpreted as the relative increase in the hazard of graft failure based on the scores of the last biopsy. As shown in Table 4, significant associations were seen for ci and ct (i.e. interstitial fibrosis and tubular atrophy, IF/TA) and cortical iptcr. Cortical iptcr was the histologic score most strongly associated with the hazard of graft failure (hazard ratio per one standard deviation unit increase in iptcr score [i.e. 0.13] above the cut-off value of iptcr = 0.10: 1.81 [95% confidence interval: 1.07–3.04; $P = 0.026$]). The hazard ratio estimate did not change substantially after adjusting for IF/TA and for time from transplantation, recipient's age, gender and Africa-American race (data not shown).

HLA-DSA strength

The results of ordinal logistic regression analysis examining the relationship between HLA-DSA strength and histologic scores are reported in Table 5. Odds ratios are interpreted as

Table 5. Association between histologic score and donor-specific antibodies strength.

Scores	Beta coefficient	OR	95% CI	*P* value
Banff scores				
g	0.36	1.43	0.76–2.68	0.27
ptc	0.74	**2.09**	**1.28–3.40**	**0.003**
i	0.44	1.55	0.89–2.56	0.12
t	0.05	1.06	0.60–2.02	0.75
v	0.36	**1.43**	**1.08–1.88**	**0.01**
ah	−0.21	0.81	0.52–1.27	0.36
cg	0.01	1.00	0.66–1.50	0.39
ci	−0.31	0.73	0.35–1.51	0.40
ct	−0.34	0.71	0.35–1.46	0.35
cv	−0.12	0.88	0.55–1.44	0.62
mm	−0.21	0.81	0.52–1.26	0.35
ti	0.12	1.16	0.59–2.29	0.67
C4d	0.34	1.41	0.83–2.37	0.21
Quantitative MVI scores				
gml	0.34	1.41	0.85–2.33	0.19
gmaxl	0.44	1.56	0.93–2.63	0.09
Cortical ptcml	0.72	**2.05**	**1.29–3.26**	**0.001**
Cortical ptcmaxl	0.75	**2.12**	**1.15–3.93**	**0.02**
Cortical iptcr	0.88	**2.42**	**1.31–4.47**	**0.005**

95% CI, 95 percent confidence interval; ah, arteriolar hyalinosis; Beta coefficient, log(OR); cg, chronic glomerulopathy; ci, interstitial fibrosis; ct, tubular atrophy; cv, chronic arteriopathy; g, glomerulitis; gmaxl, max glomerular leukocyte; gml, mean glomerular leukocytes; i, interstitial inflammation; iptcr, inflamed peritubular capillaries ratio; mm, mesangial expansion; MVI, microvascular inflammation; OR, odds ratio; ptcmaxl, max peritubular capillary leukocytes; ptcml, mean peritubular capillary leukocytes; ptc, peritubular capillaritis; ti, total inflammation; t, tubulitis; v, endoarteritis. Bold indicates statistically significant.

Results of ordinal logistic regression models examining the cross-sectional relation between each histologic scores and DSA strength at the time of biopsy. The odds ratios represent the relative increase in the probability (odds) of belonging to the higher antibody strength category (Negative, Luminex, Flow-crossmatch, CDC-crossmatch) per one standard deviation unit increase in the histologic score. Beta coefficients (log OR) are reported along with ORs in order to make easier the direct comparison between each histologic score: the larger the absolute value of the beta coefficient the stronger is the relationship between the histologic score and the antibody strength. In fact, because of missing data, number of measurements differed between histologic scores (median 65 [range 50–66]).

the relative increase in the probability (odds) of belonging to the higher antibody strength category (Luminex, Flow-XM or CDC-XM positive HLA-DSA) compared to the lower category. Findings are expressed per one standard deviation unit increase in the histologic scores. As shown in Table 5, significant correlations were shown by ptc (odds ratio [OR] 2.09, $P = 0.003$), cortical iptcr (OR: 2.42, $P = 0.005$), cortical ptcml (OR: 2.05, $P = 0.002$) and cortical ptcmaxl (OR: 2.12, $P = 0.017$). In order to provide a visual appraisal of the iptcr data, we plotted in Fig. 4 the probability of being in a specific DSA strength class based on the cortical iptcr. For instance, when iptcr is close to 0, three out of four patients will be negative for DSA; as the ratio of inflamed peritubular capillaries increases, the chance to have DSA (and AMR) will be higher: with an iptcr of approximately 0.15 (15% of inflamed peritubular capillaries), 50% of patients will show DSA and with iptcr = 0.25, 75% of patients will show DSA. A level of 50% of peritubular capillaries inflammation was rarely reached in our cohort: at this point the chance to find DSA is approximately 90%.

Correlation between cortical and medullary peritubular capillaries quantitative scores

Eleven biopsies showed both cortical and medullary tissue, and discrete regional evaluation was therefore possible. As shown in Fig. 5, there was a reasonable concordance of peritubular capillaries inflammation scores (iptcr) between the cortex and medulla (Lin's concordance correlation coefficient: 0.66, $P = 0.013$).

Reproducibility studies of Banff ptc score

To assess the impact of CD34–CD45 IHC compared to conventional stains on inter-pathologist agreement of MVI, four pathologists scored ptc score and assessed

Transplant International 2019; 32: 300–312

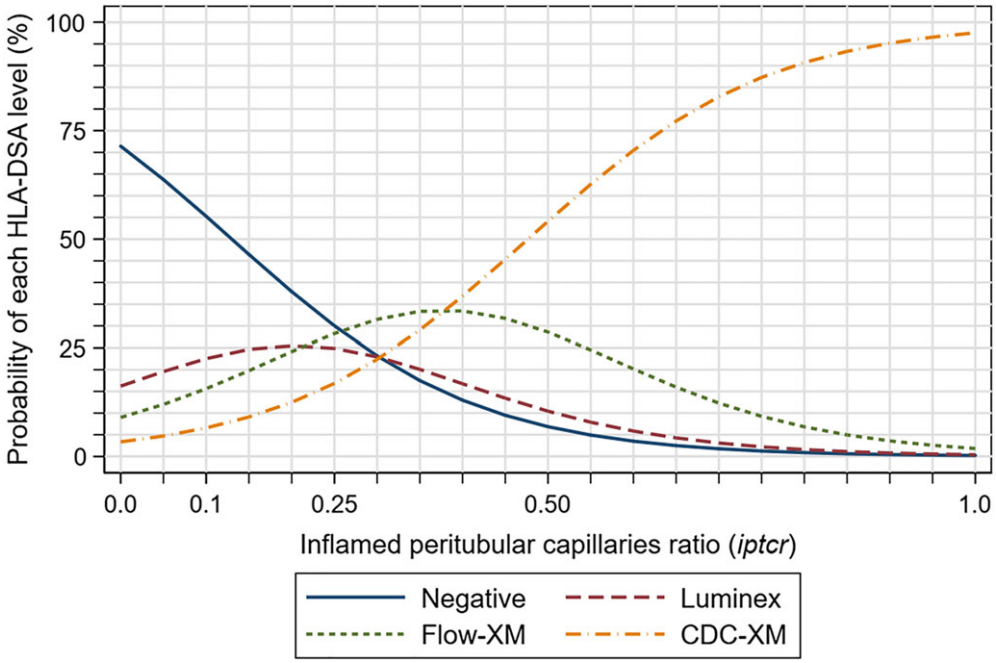

Figure 4 The plot shows the predicted probability for a patient to have negative DSA (solid blue line), positive Luminex DSA (dashed red line), positive flow-cytometry crossmatch DSA (short-dashed green line) and positive complement-dependent cytotoxic crossmatch DSA (dash-dotted yellow line) based on the measured inflamed peritubular capillaries ratio (*iptcr*) on kidney biopsy evaluated using CD34–CD45 immunohistochemical stain. Plotted predicted probabilities were calculated from the fitted ordered logistic regression model (see text). For instance, the model predicts that patients with *iptcr* above 0.50 have more than 50% probability of having circulating CDC-XM DSA, and that patients with *iptcr* below 0.10 have more than 50% probability of negative DSA, and less than 20% probability of positive Luminex DSA. 69 samples were available for serological studies at the time of the biopsy. ATR1-antibody positive and ABO-incompatible-only patients were not included. Eventually, 56 measurements were available for the analyses. Only one biopsy showed inflammatory cells in more than 50% of the peritubular capillaries. CDC-XM, complement-dependent cytotoxic crossmatch; DSA, donor-specific anti-HLA antibodies; Flow-XM, flow cytometry crossmatch.

peritubular capillaritis extent (focal versus diffuse). The use of CD34–CD45 IHC improved inter-pathologist agreement on the semiquantitative assessment of MVI: with conventional stains, inter-pathologist agreement for ptc Banff score and ptc extent was only moderate (percent agreement 0.49 and 0.54, respectively; weighted κ_G: 0.53 [95% confidence interval: 0.41–0.64] and 0.46 [0.35–0.58], respectively). By contrast, with the use of CD34–CD45 IHC, agreement improved for both Banff ptc and ptc extent (percent agreement: 0.66 and 0.65, respectively; weighted κ_G: 0.80 [0.72–0.89]) and 0.68 [0.57–0.80], respectively; $P < 0.001$ compared to conventional stain). Calculations based on unweighted kappa coefficient, and based on alternative kappa coefficient estimators yielded similar findings (see Appendix S1). As shown in Fig. 6, pathologists provided identical score for Banff ptc in two-thirds of the cases with CD34–CD45 IHC compared to in less than half of the cases with conventional stain. The improvement in the estimate of ptc extent with CD34–CD45 IHC compared to conventional stain was similar although slightly less marked (Fig. 7). The most serious disagreements (i.e., ptc 0 versus 3, diffuse ptc versus negative ptc) were rare with CD34–CD45 IHC compared to conventional stain for both ptc score and ptc extent (0.9% vs. 3% and 2.1% vs. 7.5%, respectively).

Discussion

This study demonstrates that the use of a quantitative peritubular capillaritis score, reflecting MVI in the renal cortex or medulla, predicts HLA-DSA strength at the time of biopsy and subsequent allograft loss. It also implements a new IHC stain for CD34–CD45 that substantially increases inter-pathologist agreement of the semiquantitative Banff score for peritubular capillaritis.

Because the CD34–CD45 IHC allowed precise quantification of the presence of leukocytes in peritubular capillaries and glomeruli, we developed quantitative glomerulitis and peritubular capillaritis scores based on leukocyte enumeration on digitally acquired slides. In contrast with standard Banff Classification criteria (ptc scored only in "non-scarred cortical parenchyma", excluding subcapsular, perivascular and cortico-medullary junction) [23], we performed the quantitative

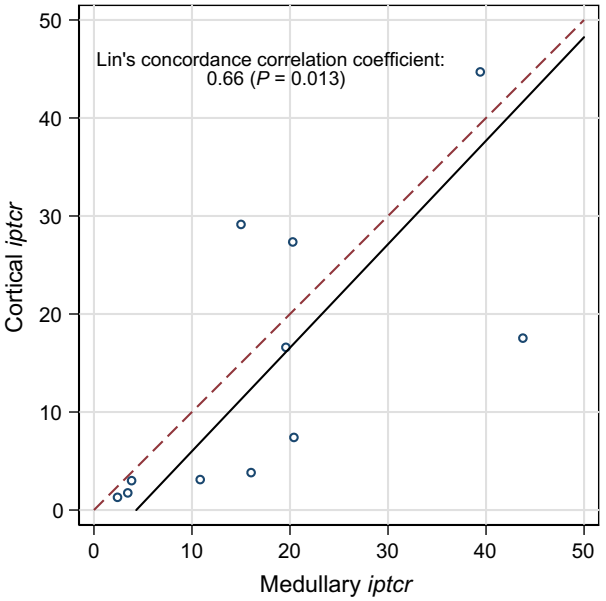

Figure 5 Scatter plot comparing the concordance between cortical (*y*-axes) and medullary (*x*-axes) inflamed peritubular capillaries ratio (*iptcr*) score. The Lin's concordance correlation coefficient was 0.66 ($P = 0.013$) indicating a reasonable concordance between cortical and medullary *iptcr*. The Lin's concordance correlation coefficient combines measures of both precision and accuracy to determine how far the observed data deviate from the dotted red line which represents the line of perfect concordance. Lin's coefficient increases in value as a function of the nearness of the data's reduced major axis (solid black line) to the line of perfect concordance (the accuracy of the data) and of the tightness of the data (blue circles) about its reduced major axis (the precision of the data).

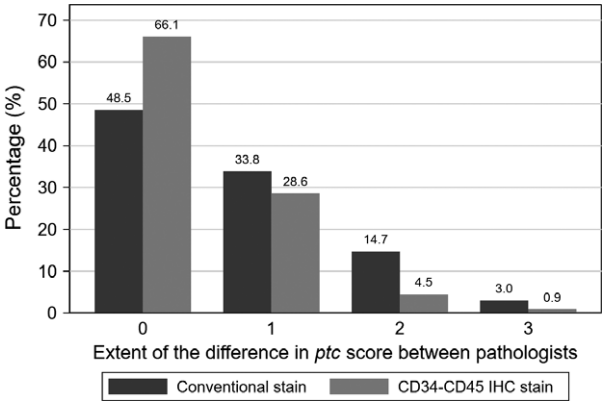

Figure 6 Differences in ptc scores between pathologists. Bars represent the observed percentage of each possible arithmetical difference in the Banff ptc score between pathologists. Dark grey bars represent differences found using conventional stain, light grey bars represent differences found using CD34–CD45 immunohistochemical (IHC) stain. Compared to conventional stain, CD34–CD45 IHC stain increased the percentage of identical evaluation (difference = 0) and decreased all other disagreement differences (difference = 1, 2 and 3, respectively). The overall difference of inter-pathologist agreement between conventional and CD34–CD45 IHC stain was statistically significant ($P < 0.001$).

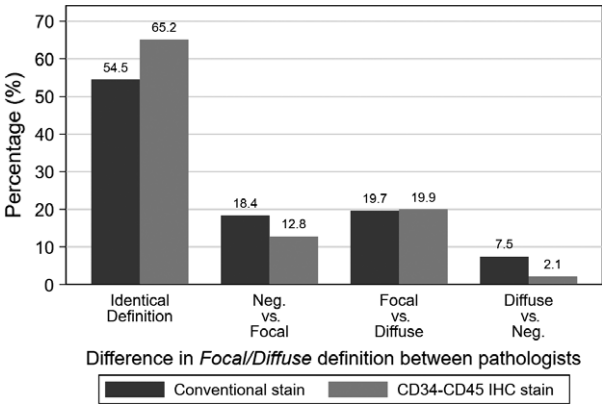

Figure 7 Differences in Focal/Diffuse definition between pathologists. Bars represent the observed percentage of each possible difference in diagnosis between pathologists. Dark grey bars represent differences found using conventional stain, light grey bars represent differences found using CD34–CD45 immunohistochemical (IHC). The use of CD34–CD45 IHC stain increased the percentage of identical evaluation and decreased the percentage of Negative Vs Focal and Diffuse Vs negative disagreements. The percentage of Focal Vs Diffuse disagreement remained unchanged. The overall difference of inter-pathologist agreement between conventional and CD34–CD45 IHC was statistically significant ($P < 0.001$).

peritubular capillaritis study on all available parenchyma. In a subset of our study population, we showed that cortical and medullary inflamed peritubular capillary ratio are significantly correlated. Interestingly, Halloran and colleagues recently showed how a molecular signature of rejection allows assessment of rejection in renal medulla [24]. Although we have no transcriptomic data in our cohort, we contend that the renal medulla histologic parameter (iptcr) may be significantly associated with HLA-DSA presence and allograft survival. Larger cohorts of biopsies complete of cortex and medulla are needed to confirm our findings.

Given the strong correlation between cortical iptcr and HLA-DSA strength, which was supported by the relationship between iptcr and graft outcome, we developed a prognostic model for HLA-DSA strength at the time of biopsy based on the percentage of cortical peritubular capillaries with one or more luminal leukocytes (shown as a nomogram in Fig. 4). To our knowledge, this is the first evidence of a relationship between HLA-DSA strength and MVI intensity. If confirmed, it will provide new opportunities for defining clinically relevant HLA-DSA, for harmonizing antibody data across centers, and for enabling rapid clinical decision-making when HLA-DSA results may not yet be available. We acknowledge that the predictive model was based on a small sample of data, and that our findings should be validated in further studies. However, the

cohort included high immunological risk patients with HLA-DSA detected across a broad range of strength levels, representing a powerful setting to study MVI and its clinical correlations. Indeed, we deliberately included ABO- and HLA-incompatible transplants, of whom almost 32% had at least one previous kidney transplant.

Our results are in line with observations of Viglietti and colleagues who recently developed a prognostic score for patients with AMR: the score includes eGFR at the time of diagnosis, presence of chronic allograft glomerulopathy, IF/TA degree, *de novo* HLA-DSA and intensity of peritubular capillaritis; the authors suggest that this score should drive therapeutic approach to patients with AMR [25]. Our iptcr score alone can predict DSA strength as well as graft survival, an improvement on the ptc score used by Viglietti.

Reproducibility studies of Banff scoring system showed less than robust inter-pathologist agreement [4–8]. Furness *et al.* [4,5] analyzed numerous parameters potentially influencing reproducibility of Banff classification. Slides circulated among European transplant pathologists showed alarmingly low inter-observer reproducibility. While ptc was not formally included in the Banff classification at the time of this study, investigators showed that peritubular capillary neutrophils were essentially not reproducible (κ value 0.13) [4]. In a pilot study involving three pathologists, Gibson *et al.* [9] demonstrated fair-moderate intra-observer reproducibility for the ptc score (mean κ: 0.62) and weaker inter-pathologist reproducibility (mean κ: 0.49); in a larger study with six pathologists, a simplified ptc-scoring system showed fair-moderate, albeit overall weaker, inter-pathologist reproducibility (κ-values: 0.32–0.43). In our study, the improved reproducibility (κ-coefficient from 0.52–0.80 and 0.46–0.68 for ptc score and ptc extension, respectively, $P < 0.001$) might be related to the fact that CD34–CD45 IHC highlights all peritubular capillaries, providing a better contrast to distinguish intracapillary versus interstitial leukocytes and more precise enumeration of intracapillary leukocytes.

A limitation for our approach is the complexity of quantitative MVI assessment: a digital slide and an appropriate software are necessary, and meticulous cell-counting is time-intensive. Additional studies are needed to establish if quantification can be assessed on limited areas of renal tissue.

In conclusion, we have shown that CD34–CD45 IHC allows more reproducible assessment of Banff ptc compared to conventional stain, and that CD34–CD45-based quantitative assessment of MVI may predict HLA-DSA strength at the time biopsy. Future studies are needed to standardize CD34–CD45 IHC stain for routine clinical use, and to validate our findings on the predictive value on the quantitative assessment of MVI in additional cohorts and clinical centers.

Authorship

MD: collected data, reviewed renal biopsies for reproducibility studies and drafted the manuscript. UM: analyzed data and drafted the manuscript. AZR: formulated the initial hypothesis, designed the study, reviewed renal biopsies for reproducibility studies, supervised the research and helped draft the manuscript. JL: collected data. DEK and SMB: reviewed renal biopsies for reproducibility studies. AMJ: provided DSA data and drafted the manuscript. LJA, SMH and NCM: assessed the results and gave critical feedback on the research and design of the study. All authors: read and approved the final draft of the manuscript.

Funding

The authors have declared no funding.

Conflicts of interest

The authors of this manuscript have no conflicts of interest to disclose as described by Transplant International.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Supplementary methods.

Table S1. Distribution of Banff scores.

Table S2. (A) Percent agreement and different unweighted kappa coefficients on ptc with CD34–CD45 IHC stain. (B) Percent agreement and different unweighted kappa coefficients on ptc with conventional stain. (C) Percent agreement and different weighted kappa coefficients on ptc with CD34–CD45 IHC stain. (D) Percent agreement and different weighted kappa coefficients on ptc with conventional stain.

Table S3. (A) Percent agreement and different unweighted kappa coefficients on ptc extension (focal versus diffuse) with CD34–CD45 IHC stain. (B) Percent agreement and different unweighted kappa coefficients on ptc extension (focal versus diffuse) with conventional

stain. (C) Percent agreement and different weighted kappa coefficients on ptc extension (focal versus diffuse) with CD34–CD45 IHC stain. (D) Percent agreement and different weighted kappa coefficients on ptc extension (focal versus diffuse) with conventional stain.

Table S4. (A) Differences of ptc percent agreement and of unweighted kappa coefficients between CD34–CD45 IHC stain and conventional stain. (B) Differences of ptc percent agreement and of weighted kappa coefficients between CD34–CD45 IHC stain and conventional stain. (C) Differences of ptc extention (focal versus diffuse) agreement and of unweighted kappa coefficients between conventional stain and CD34–CD45 IHC stain. (D) Differences of ptc extention (focal versus diffuse) agreement and of weighted kappa coefficients between conventional stain and CD34–CD45 IHC stain.

Figure S1. Change in ptc score grading with use of CD34–CD45 IHC stain (*y* axis) compared to conventional stain (*x* axis).

Figure S2. Change in ptc extension grading with use of CD34–CD45 IHC stain (*y* axis) compared to conventional stain (*x* axis).

REFERENCES

1. Loupy A, Lefaucheur C, Vernerey D, *et al.* Complement-binding anti-HLA antibodies and kidney-allograft survival. *N Engl J Med* 2013; **369**: 1215.
2. Michielsen LA, van Zuilen AD, Krebber MM, Verhaar MC, Otten HG. Clinical value of non-HLA antibodies in kidney transplantation: still an enigma? *Transplant Rev (Orlando)* 2016; **30**: 195.
3. Colvin RB, Smith RN. Antibody-mediated organ-allograft rejection. *Nat Rev Immunol* 2005; **5**: 807.
4. Furness PN, Taub N, Project CoERT-PAPC. International variation in the interpretation of renal transplant biopsies: report of the CERTPAP Project. *Kidney Int* 2001;**60**:1998.
5. Furness PN, Taub N, Assmann KJ, *et al.* International variation in histologic grading is large, and persistent feedback does not improve reproducibility. *Am J Surg Pathol* 2003; **27**: 805.
6. Marcussen N, Olsen TS, Benediktsson H, Racusen L, Solez K. Reproducibility of the Banff classification of renal allograft pathology. Inter- and intraobserver variation. *Transplantation* 1995; **60**: 1083.
7. Solez K, Hansen HE, Kornerup HJ, *et al.* Clinical validation and reproducibility of the Banff schema for renal allograft pathology. *Transplant Proc* 1995; **27**: 1009.
8. Gough J, Rush D, Jeffery J, *et al.* Reproducibility of the Banff schema in reporting protocol biopsies of stable renal allografts. *Nephrol Dial Transplant* 2002; **17**: 1081.
9. Gibson IW, Gwinner W, Bröcker V, *et al.* Peritubular capillaritis in renal allografts: prevalence, scoring system, reproducibility and clinicopathological correlates. *Am J Transplant* 2008; **8**: 819.
10. Loupy A, Hill GS, Suberbielle C, *et al.* Significance of C4d Banff scores in early protocol biopsies of kidney transplant recipients with preformed donor-specific antibodies (DSA). *Am J Transplant* 2011; **11**: 56.
11. de Kort H, Willicombe M, Brookes P, *et al.* Microcirculation inflammation associates with outcome in renal transplant patients with de novo donor-specific antibodies. *Am J Transplant* 2013; **13**: 485.
12. Haas M, Sis B, Racusen LC, *et al.* Banff 2013 meeting report: inclusion of c4d-negative antibody-mediated rejection and antibody-associated arterial lesions. *Am J Transplant* 2014; **14**: 272.
13. Pusztaszeri MP, Seelentag W, Bosman FT. Immunohistochemical expression of endothelial markers CD31, CD34, von Willebrand factor, and Fli-1 in normal human tissues. *J Histochem Cytochem* 2006; **54**: 385.
14. Solez K, Colvin RB, Racusen LC, *et al.* Banff 07 classification of renal allograft pathology: updates and future directions. *Am J Transplant* 2008; **8**: 753.
15. Zachary AA, Lucas DP, Detrick B, Leffell MS. Naturally occurring interference in Luminex assays for HLA-specific antibodies: characteristics and resolution. *Hum Immunol* 2009; **70**: 496.
16. Liao JJ, Lewis JW. A note on concordance correlation coefficient. *PD'A J Pharm Sci Technol* 2000; **54**: 23.
17. Orandi BJ, Luo X, Massie AB, *et al.* Survival benefit with kidney transplants from HLA-incompatible live donors. *N Engl J Med* 2016; **374**: 940.
18. Montgomery RA, Lonze BE, King KE, *et al.* Desensitization in HLA-incompatible kidney recipients and survival. *N Engl J Med* 2011; **365**: 318.
19. Philogene MC, Bagnasco S, Kraus ES, *et al.* Anti-angiotensin II type 1 receptor and anti-endothelial cell antibodies: a cross-sectional analysis of pathological findings in allograft biopsies. *Transplantation* 2017; **101**: 608.
20. Gwet KL. *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Multiple Raters*, 4th edn. Gaithersburg, MD, USA: Advanced Analytics, LLC, 2014.
21. Gwet KL. Testing the difference of correlated agreement coefficients for statistical significance. *Educ Psychol Meas* 2016; **76**: 609.
22. Williams RL. A note on robust variance estimation for cluster-correlated data. *Biometrics* 2000; **56**: 645.
23. Racusen LC, Solez K, Colvin RB, *et al.* The Banff 97 working classification of renal allograft pathology. *Kidney Int* 1999; **55**: 713.
24. Madill-Thomsen KS, Wiggins RC, Eskandary F, Böhmig GA, Halloran PF. The effect of cortex/medulla proportions on molecular diagnoses in kidney transplant biopsies: rejection and injury can be assessed in medulla. *Am J Transplant* 2017; **17**: 2117.
25. Viglietti D, Loupy A, Aubert O, *et al.* Dynamic prognostic score to predict kidney allograft survival in patients with antibody-mediated rejection. *J Am Soc Nephrol* 2017; **29**: 606.