OPEN

Noninvasive Diagnosis of Acute Rejection in Renal Transplant Patients Using Mass Spectrometric Analysis of Urine Samples: A Multicenter Diagnostic Phase III Trial

Wilfried Gwinner, MD,[1] Annika Karch, PhD,[2] Jan H. Braesen, MD,[3] Abedalrazag A. Khalifa, MD,[3] Jochen Metzger, PhD,[4] Maarten Naesens, MD,[5,6] Elisabet Van Loon, MD,[5,6] Dany Anglicheau, MD,[7] Pierre Marquet, PhD,[8] Klemens Budde, MD,[9] Mareen Matz, PhD,[9] Wolfgang Arns, MD,[10] Michael Fischereder, MD,[11] Antje Habicht, MD,[11] Ute Eisenberger, MD,[12] Anja Mühlfeld, MD,[13] Martin Busch, MD,[14] Michael Wiesener, MD,[15] Irina Scheffner, PhD,[1] and Armin Koch, PhD[2]

Background. Timely recognition and treatment of acute kidney graft rejection is important to prevent premature graft failure. A predefined urinary marker set for acute T cell–mediated rejection (TCMR) containing 14 peptides was tested for this purpose in a multicenter in-place validation study. **Methods.** Three hundred twenty-nine prospectively collected and 306 archived urine samples from 11 transplant centers in Germany, France, and Belgium were examined. Samples were taken immediately before a biopsy, performed for graft dysfunction within the first transplant year. Primary outcomes were sensitivity and specificity of the marker set for the diagnosis of biopsy-proven acute TCMR, with prespecified thresholds of 83% for sensitivity and 70% for specificity. **Results.** Eighty-two patients (13%) had acute TCMR grade I–III. In relation to the biopsy diagnosis of TCMR, the sensitivity of the urine test was 0.66 (95% confidence interval, 0.56-0.76) and the specificity 0.47 (95% confidence interval, 0.43-0.51), with an area under the curve (AUC) of 0.60. The different TCMR grades I–III were not reflected by the marker set, and borderline TCMR was not specifically detected. Secondary independent masked assessment of biopsies consented by 2 pathologists revealed an interobserver kappa value of 0.49 for diagnosing TCMR, compared with the local center's diagnosis. Using this consensus diagnosis, the AUC of the urine test was 0.63 (sensitivity 0.73, specificity 0.45). Post hoc optimization of the marker set improved the diagnostic performance in the study cohort (AUC 0.67) and in an independent patient cohort (AUC 0.69). **Conclusions.** This study illustrates the difficulty of proteomics-based diagnosis of TCMR and highlights the need for rigorous independent in-place validation and optimization of diagnostic biomarkers.

(*Transplantation Direct* 2022;8: e1316; doi: 10.1097/TXD.0000000000001316).

Received 3 February 2022.

Accepted 14 February 2022.

[1] *Department of Nephrology, Hannover Medical School, Hannover, Germany.*

[2] *Institute for Biostatistics, Hannover Medical School, Hannover, Germany.*

[3] *Department of Pathology, Hannover Medical School, Hannover, Germany.*

[4] *Mosaiques Diagnostics GmbH, Hannover, Germany.*

[5] *Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium.*

[6] *Department of Nephrology, UZ Leuven, Leuven, Belgium.*

[7] *Department of Nephrology and Kidney Transplantation, Necker Hospital, AP-HP and Inserm U1151, Paris, France.*

[8] *Pharmacology and Transplantation, INSERM U1248, Université de Limoges, CHU de Limoges, Limoges, France.*

[9] *Charité Universitätsmedizin Berlin, Berlin, Germany.*

[10] *Cologne Merheim Medical Center, Cologne, Germany.*

[11] *LMU-Klinikum, Munich, Germany.*

[12] *University Hospital Essen, Duisburg-Essen, Essen, Germany.*

[13] *Clinic for Renal and Hypertensive Disorders, Rheumatological and Immunological Diseases, University Hospital Aachen, Aachen, Germany.*

[14] *Department of Internal Medicine III/Collaborative Kidney Transplant Center, University Hospital Jena—Friedrich Schiller University, Jena, Germany.*

[15] *Department of Nephrology and Hypertension, Universitätsklinikum Erlangen, University of Erlangen-Nürnberg, Erlangen, Germany.*

This study was supported by a grant from the Deutsche Forschungsgemeinschaft (GW 4/6-1). The sequence analysis of peptides was supported within the ROCKET study, funded by the Bundesministerium für Bildung und Forschung (BMBF) under the frame of ERACoSysMed-2, the European Research Area Network for Systems Medicine in Clinical Research and Medical Practice (grant reference no. JTC2_29). The external validation cohort was provided by the BIOMARGIN study, funded by the Seventh Framework Programme (FP7) of the European Commission, in the HEALTH.2012.1.4-1 theme of "innovative approaches to solid organ transplantation" (grant reference no. 305499).

J.M. is an employee of Mosaiques Diagnostics GmbH. The other authors declare no conflicts of interest.

W.G., A.Ka., and A.Ko. participated in research design. WG., A.Ka., A.Ko., J.M., and I.S. participated in data analysis. All authors were involved in manuscript writing/revision and performance of the research/study.

Supplemental digital content (SDC) is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (www.transplantationdirect.com).

Correspondence: Wilfried Gwinner, MD, Department of Nephrology, Hannover Medical School, Carl-Neuberg-Str.1, D-30625 Hannover, Germany. (gwinner.wilfried@mh-hannover.de).

ISSN: 2373-8731

DOI: 10.1097/TXD.0000000000001316

Introduction

Acute rejections have an unfavorable prognosis on long-term kidney graft survival.[1,2] Timely detection and appropriate treatment of rejection is important to prevent irreversible tissue damage and decreasing graft function. However, current practice with regular monitoring of graft function by serum creatinine and performing a graft biopsy upon functional impairment alone may be insufficient.[3]

To improve early detection of rejection, many noninvasive tests in blood and urine have been explored in the past, mostly using omics approaches at different molecular levels. However, none of these tests are established in clinical practice yet, mainly because of lacking sensitivity and specificity, and insufficient validation by appropriate independent studies.[3-5]

Based on a previously established urinary peptide marker set for acute T cell–mediated rejection (TCMR),[6] this study was conducted to examine the suitability of the marker set under clinical practice conditions. The study was planned as a prospective multicenter diagnostic phase III study.[7] Because of insufficient recruitment, the study was amended to include additional archived samples collected using the same sampling protocol.

The primary objective was to demonstrate that the urine marker test has sufficient accuracy in detecting acute TCMR, compared to the reference standard "biopsy diagnosis." Accordingly, primary outcomes were sensitivity and specificity of the marker set for the diagnosis of TCMR, using prespecified thresholds for these parameters. Secondary endpoints were to examine the marker set in relation to different severity grades of rejection and to determine limitations of the test in terms of confounding factors that influence its accuracy.[7]

MATERIALS AND METHODS

Study Design

The study (NCT01315067) was designed as a prospective, single-arm, multicenter, phase III diagnostic study of the urinary peptide marker set ("test") to examine its diagnostic performance in relation to the diagnosis by a graft biopsy ("reference standard"). Details on the study protocol were reported previously.[7] Study procedures complied with the Declarations of Helsinki and Istanbul,[8,9] were approved by the Ethics Committee of the Hannover Medical School, and complied with the local regulations of the participating centers. Written informed consent was obtained from each patient.

From patients who were planned to have a kidney graft biopsy within the first year after transplantation for unexplained graft dysfunction according to the medical judgment of the local center, a spot urine sample was obtained immediately before biopsy and frozen at –20 °C. Urine protein determination and sediment analysis were performed in parallel. Recipient and donor data were entered into electronic Case Report Forms build with SecuTrial for central monitoring, including the estimated glomerular filtration rate (eGFR) at biopsy, baseline serum creatinine before graft impairment and factors potentially acting as confounder of the urine test (eg, cytomegalovirus, BK virus and urinary tract infection, diabetes, hypertension, medication, graft hydronephrosis and artery stenosis, and delayed graft function). Graft function and further rejection episodes within 6 mo after the index biopsy were recorded for secondary analyses.

Participants and Study Centers

Adult kidney and combined kidney/pancreas transplant recipients were included with 1 sample and biopsy/patient. Patient enrollment started in October 2011 and ended in January 2016. Because of insufficient recruitment, the study was amended in October 2015 to include additional archived samples collected using the same sampling protocol. Patients (n = 329) were prospectively recruited in the German transplant centers Hannover, Aachen, Essen, Freiburg, Cologne, Jena, Munich, Erlangen, and Berlin. Archived samples (n = 306, collected in 2008–2011) were from the centers Necker Hospital in Paris (France), Leuven (Belgium), Hannover, and Berlin (**Table S1, SDC,** http://links.lww.com/TXD/A411).

Laboratory Examination and Clinical Variable Definitions

The analysis of urine peptides by capillary electrophoresis coupled to mass spectrometry (MS) was performed as described previously in detail.[6] Classification of samples into the categories "rejection" and "no rejection" was based on the predefined peptide pattern for TCMR and a cut-point of –0.25 established in the preceding study, which also describes details on marker selection and establishment of the marker model using a support vector machine approach. The marker set was composed of different collagen chain fragments, mainly type 1 alpha, subtypes I–IV,[6,7] normalized to a high-abundance peptide set normally found in urine of healthy individuals and patients with renal disease.[10] Routine measures to ensure integrity of samples and reliability of procedures are detailed elsewhere.[11] Test reproducibility was evaluated by repeated determination of 69 patient samples of this study. The Krippendorff's alpha value for these probe pairs was 0.80 (95% confidence interval [CI], 0.69-0.86).

Peptide sequences for the Capillary electrophoresis (CE)–MS peptide marker were derived from Mosaiques peptide database,[12] which contains sequence information from analysis of human urine samples on a Dionex Ultimate 3000 RSLS nanoflow system (Dionex, Camberley, UK) and a Beckman CE/Orbitrap Q Exactive plus combination (Thermo Scientific, Waltham, MA).[13] Spectra files were processed with Proteome Discoverer 2.4 (Thermo Scientific), setting the precursor mass tolerance to 5 ppm and the fragment mass tolerance to 0.05 Da. This was followed by a SEQUEST search against the UniProt human nonredundant database (https://www.uniprot.org/) without any protease specificity or fixed modification, but considering oxidation of methionine, lysine, and proline as variable modifications. Only high confidence sequences with an Xcorr-value ≥ 1.9 without unmodified cysteine (due to nonreducing conditions) were accepted.[14] The strong correlation of the peptide's mobility in CE at the operating pH of 2 with its number of basic amino acids served as another selection criterion to avoid false sequence assignments.[15]

Biopsies were evaluated according to the Banff 2013 classification[16,17] by pathologists of the local centers. BK nephropathy was diagnosed by histochemical detection of the SV40 antigen. A subset of 409 biopsies was re-evaluated centrally by 2 pathologists (J.H.B. and A.K.) who were masked to the results of the original biopsy assessment and urine test and who agreed by consensus on a diagnosis. These results were used in a secondary analysis of the test performance.

The eGFR was calculated with the Cockcroft–Gault formula (mL/min/1.7 m^2). Urinary tract infection was defined by leukocyturia (dipstick positive and/or 5–10 leukocytes/microscopic field in a urinary sediment) in combination with a urine culture with $>10^4$ bacterial colonies, with or without clinical symptoms.[18] Viral infections (hepatitis B/C, BK virus, and

cytomegalovirus) were determined by locally available nucleic acid tests, preformed antibodies by the lymphocytotoxic panel reactive antibody test. Delayed graft function was defined as <500 mL urine within 24 h posttransplantation and/or need of dialysis within the first week (excluding cases with dialysis solely because of hyperkaliemia). Because this was a diagnostic study, rejections were treated according to the judgment and protocols of the participating centers.

Statistics

Endpoints

Based on the primary objective to demonstrate sufficient accuracy of the urine test to detect acute TCMR compared to the reference "biopsy diagnosis," sensitivity and specificity of the marker set for the diagnosis of TCMR were calculated as point estimators with 2-sided 95%-Wald CIs. Prespecified thresholds for rejecting the null hypothesis were a lower bound of the corresponding CI >83% for sensitivity and >70% for specificity. Secondary objectives included examination of the marker set in relation to the severity grades of TCMR, sensitivity, specificity, and predictive values for subgroups regarding center-specific results, infections, delayed graft function, other concomitant biopsy findings, and medication.

Sample Size

Separate calculations for sensitivity and specificity were based on a presumed TCMR prevalence of 25% and the estimators derived from the previous phase II study with 91% for sensitivity and 76% for specificity. A necessary number of 150 TCMR cases and 440 cases without TCMR was determined.[7] Therefore, recruitment of 600 patients was planned.

Statistical Analyses

Descriptive analyses were performed in the whole cohort of 635 patients and separately for prospectively collected and archived samples (Figure 1). Applying the "intention to treat (diagnose)" principle, the primary analysis included all patients fulfilling the study inclusion/exclusion criteria and having a conclusive biopsy result (n = 629). Missing urine test values were imputed as having the incorrect diagnosis. A sensitivity analysis of the primary and all further efficacy analyses was performed in the per-protocol dataset, that is, patients with available conclusive biopsy and urine test result (n = 624).

Continuous variables are given as means ± SD or medians with 25/75% quartiles and compared with the *t*-test or Wilcoxon test. Categorical data were analyzed with the chi-square test. Receiver-operating characteristics with the area under the curve (AUC), specificity, sensitivity, positive and negative predictive values, and their 95% CIs describe the urine test performance. Kappa statistics describe the results of the primary and secondary biopsy assessment. Missing values >n = 15 are reported in the tables. Statistical significance is assumed for $P < 0.05$ (2-tailed).

RESULTS

Characteristics of Patients at Transplantation and at the Index Biopsy

Pretransplant and transplant data are reported in Table 1. Causes of end-stage renal failure were more often biopsy-confirmed in prospectively recruited patients. Preformed panel reactive antibodies, living donor, and AB0 blood group-incompatible transplantations were more prevalent in these patients, along with more frequent use of peritransplant plasmapheresis. Induction therapy with antilymphocyte globulins or immune globulin G was more frequently given to patients with archived samples. Cold ischemia times were longer in these patients, without differences in delayed graft function compared with the prospectively recruited patients.

At the time of the index biopsy (Table 2), median serum creatinine concentration was 210 µmol/L, corresponding to an eGFR of 37 ± 18 mL/min, without relevant difference between patients with prospectively collected and archived samples. The lowest median serum creatinine within 30 d before the index biopsy was 178 µmol/L. Urinary tract infections and other bacterial infections were more common in patients with archived samples, which was not reflected by relevantly higher C-reactive protein or white blood cell count. Yet, this patient group received more often antibacterial treatment as depicted in **Table S2, SDC,** http://links.lww.com/TXD/A411, which also reports the immunosuppressive therapy and other medication. Complications of the biopsy occurred in 4.3% (perirenal hematoma; n = 14, hematuria; n = 9, arterio-venous fistula; n = 4).

Results on the Index Biopsies

A third of the index biopsies were performed within 14 d posttransplantation (Table 3). According to the Banff criteria, 58.4% of the biopsies were fully adequate and 18.7% minimal adequate. Inadequate biopsies were present in 22.5%, with 2.8% completely noninformative biopsies due to scarring. TCMR was diagnosed in 82 samples (12.9%) and borderline TCMR in 157 (24.7%), with minor differences between prospective and archived samples. Incidence of TCMR and borderline TCMR was highest within 14 d posttransplantation. Criteria of antibody-mediated rejection (AMR), glomerulitis, and peritubular capillaritis were more prevalent in archived samples. Insufficient data on donor-specific antibodies precluded the separation of antibody-positive and -negative cases. Other relevant biopsy findings are also shown in Table 3. A complete inventory of TCMR cases is depicted in Table 4, demonstrating a high proportion with additional glomerulitis and peritubular capillaritis in >40%, compared with <20% in borderline TCMR or without TCMR. Glomerulitis and peritubular capillaritis associated with peritubular C4d positivity (no TCMR: 36.9% versus 6.2% in cases without glomerulitis/peritubular capillaritis, borderline TCMR: 31.0% versus 7.0%, TCMR grade I: 37.5% versus 15.0%, TCMR grade II–III: 30.0% versus 15.4%; AB0-incompatible transplantations excluded).

Poorer graft function at biopsy was observed with TCMR grade II–III, compared with patients without TCMR (Figure 2A). The rise in serum creatinine at biopsy compared with the baseline value was highly variable in all groups with and without rejection (Figure 2B). Most patients with TCMR received rejection treatments, whereas borderline TCMR cases were treated less frequently (115/157). Notably, 9.5% of patients without any rejection signs received treatment (Figure 2C). Details of rejection treatments are shown in **Table S3, SDC,** http://links.lww.com/TXD/A411.

Evaluation of the Urinary Peptide Marker Set

Application of the predefined urinary peptide marker set to the 82 biopsy-confirmed samples with TCMR and 547 samples without TCMR (intention-to-treat dataset) showed

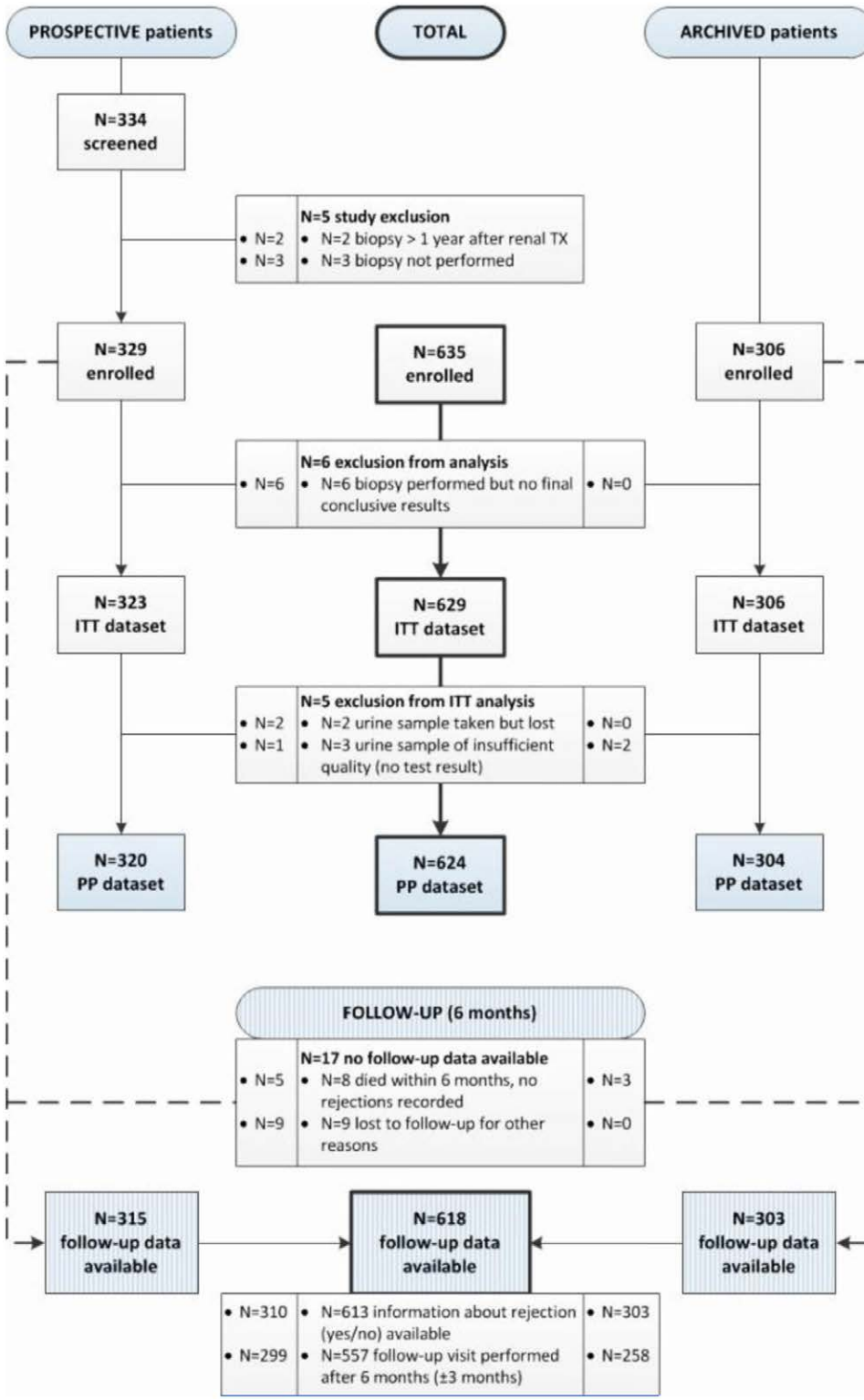


FIGURE 1. Disposition of patients of the study. The primary analysis was performed using the ITT principle containing all patients fulfilling the study inclusion/exclusion criteria and having a available conclusive biopsy result (n = 629). A sensitivity analysis of the primary analysis and all further efficacy analyses were performed in the PP dataset containing all patients fulfilling the study inclusion/exclusion criteria and having a available conclusive biopsy result and available MS urine test result (n = 624). ITT, intention to treat; MS, mass spectrometry; PP, per protocol.

a sensitivity of 0.66 (95% CI, 0.56-0.76) and a specificity of 0.47 (95% CI, 0.43-0.51) to diagnose acute TCMR, with an AUC of 0.60. Prospectively collected and archived samples had similar results (Figure 3). The median MS classifier score was −0.16 (interquartiles −0.59, 0.29) for the whole patient group (prospective samples −0.14; interquartiles −0.57, 0.34; archived samples −0.17, interquartiles −0.62, 0.27). The distribution of classifier scores among different TCMR grades

TABLE 1.

Characteristics of patients at transplantation and donor data

	Total samples n = 635	%	Prospectively recruited n = 329	%	Archived samples n = 306	%	*P*
Age (y)	53.0 ± 14.1		52.7 ± 14.3		53.3 ± 13.9		0.6234
Sex (m/f)	403/232	63.5/36.5	211/118	64.1/35.9	192/114	62.7/37.3	0.7165
Cause of end-stage renal failure[a]							
Glomerulonephritis, biopsy-proven (11–17, 19)	136	21.4	83	25.2	53	17.3	
Suspected glomerulonephritis, no biopsy (10)	22	3.5	14	4.3	8	2.6	
Interstitial nephritis (20–24, 29–31, 33, 39)	62	9.8	38	11.6	24	7.8	
Cystic kidney disease (40, 41, 43, 49)	101	15.9	52	15.8	49	16.0	
Alport's syndrome (51)	10	1.6	5	1.5	5	1.6	
Other congenital disease (50, 53, 59, 60, 61, 63, 66)	19	3.0	7	2.1	12	3.9	
Vascular diseases (70–72)	43	6.8	17	5.2	26	8.5	
Polyarteriitis, Wegener's granulomatosis (73, 74)	11	1.7	8	2.4	3	1.0	
Diabetic nephropathy (80)	67	10.6	28	8.5	39	12.8	
Other secondary systemic disease (83–88)	29	4.6	16	4.9	13	4.3	
Miscellaneous diseases (90, 92, 93, 95, 96, 99)	19	3.0	12	3.7	7	2.3	
No identified cause; etiology uncertain (00)	116	18.3	49	14.9	67	21.9	
Biopsy-proven cause of end-stage renal failure	200	31.5	138	41.9	62	20.3	<0.0001
Preemptive Tx	43	6.8	17	5.2	26	8.5	0.1100
Retransplant	96	15.1	54	16.4	42	13.8	0.3536
Combined pancreas/kidney Tx	12	1.9	9	2.7	3	1.0	0.1000
Donor age (y)	56.0 ± 14.1		56.0 ± 13.7		56.0 ± 14.6		0.9880
Donor sex (m/f/unknown)	261/359/15	41.1/56.5/24	127/193/9	38.6/58.7/2.7	134/166/6	43.8/54.2/2.0	0.2095
Deceased donor	474	74.9	230	70.1	244	80.0	0.0047[b]
Living donor (blood-related/not blood-related)	74/85	11.7/13.4	39/59	11.9/18.0	35/26	11.5/8.5	0.0346[c]
ABO blood group-incompatible living donor Tx	32	5.0	25	7.6	7	2.3	0.0414[d]
Cold ischemia time (h)	11.3 ± 7.7		9.4 ± 6.1		13.2 ± 8.6		<0.0001
Delayed graft function/unknown	134/22	21.1/3.5	66/15	20.1/4.6	68/7	22.2/2.3	0.6058
Dialysis after Tx	166	26.1	85	25.8	81	26.5	0.8349
HLA mismatch							
A (0/1/2)	150/315/155	24/51/25	102/154/71	31/47/22	48/161/84	16/55/29	<0.0001
B (0/1/2)	96/294/230	16/47/37	66/150/111	20/46/34	30/144/119	10/49/41	0.0024
DR (0/1/2)	165/314/141	27/51/23	94/155/78	29/47/24	71/159/63	24/54/22	0.2234
Unknown	15		2		13		
Panel reactive antibodies (%)							0.0002[e]
0	344	81.7	208	75.6	136	93.2	
>0–30	34	8.1	28	10.2	6	4.1	
>30–<85	34	9.1	31	11.3	3	2.1	
>85	9	2.1	8	2.9	1	0.7	
Unknown	214	33.7	54	16.4	160	52.3	
Induction therapy							<0.0001[e]
Interleukin-2 receptor antibodies	392	62	241	73	151	51	
Antilymphocyte globulins	131	21	62[f]	19	70	23	
Rituximab	34	5	18	6	16	6	
Eculizumab	4	1	3	1	1	0	
Immune globulin G	65	10	0	--	65	22	
None	26	4	21	6	5	2	
Unknown	14	2.2	2	1.0	12	3.9	
Plasmapheresis/immune adsorption peri-Tx[f]	73	12.5	51	17.8	22	7.4	0.0001
Initial immunosuppressive maintenance therapy							
Cyclosporine A	137	21.6	88	26.7	49	16.0	0.0001
Tacrolimus	474	74.7	241	73.3	233	76.1	0.2904
Mycophenolate mofetil, mycophenolic acid	598	94.2	308	93.6	290	94.8	0.2441
Sirolimus, everolimus	27	4.3	21	6.4	6	2.0	0.0058
Steroids	615	96.9	329	100	286	93.5	<0.0001

[a]EDTA code in brackets.
[b]Deceased vs living donor Tx.
[c]Blood-related vs not blood-related living donor Tx.
[d]Blood group-incompatible vs compatible living donor Tx.
[e]Results of a chi-square test over all categories between prospectively vs archived samples.
[f]Missing information in 50 patients.
Tx, transplantation.

TABLE 2.

Clinical and laboratory data at the time of the index biopsy

	Total samples n = 635	%	Prospectively recruited n = 329	%	Archived samples n = 306	%	Pvalue
Body height (cm)	172 ± 10		173 ± 10		171 ± 9[a]		0.0140
Body weight (kg)	76 ± 162		79 ± 16		73 ± 16		<0.0001
Systolic blood pressure (mm Hg)	137 ± 19		135 ± 18		140 ± 19		0.0003
Diastolic blood pressure (mm Hg)	77 ± 12		77 ± 11		77 ± 12		0.8150
Coronary heart disease (unknown)	92 (26)	14.5 (4.1)	48 (19)	14.6 (5.8)	44 (7)	14.4 (2.3)	0.4590
Heart failure							<0.0001[b]
Grade I	16	2.5	16	4.9	0	0	
Grade II	14	2.2	13	4.0	1	0.3	
Grade III–IV	7	1.1	5	1.5	2	0.7	
Unknown	30	4.7	23	7.0	7	2.3	
Diabetes type I	56	8.8	11	3.3	45	14.7	<0.0001
Diabetes type II	87	13.7	51	15.5	36	11.8	0.2039
Replicative hepatitis B (unknown)	8 (60)	1.3 (9.4)	6 (20)	1.8 (6.1)	2 (40)	0.7 (13.1)	0.2245
Replicative hepatitis C (unknown)	12 (187)	1.9 (29.4)	8 (21)	2.4 (6.4)	4 (166)	1.3 (54.2)	0.9876
Cytomegaly virus infection or viremia (unknown)	66 (78)	11.8 (12.3)	30 (0)	9.1 (–)	36 (78)	15.8 (25.5)	0.0166
BK viremia (unknown)	43 (176)	6.8 (22.7)	23 (77)	7.0 (23.4)	20 (99)	6.5 (32.4)	0.8449
Urinary tract infection (unknown)	160 (33)	25.2 (5.0)	69 (16)	21.0 (4.9)	91 (16)	29.7 (5.2)	0.0110
Other bacterial infection	42	6.7	8	2.4	34	11.3	<0.0001
C-reactive protein (mg/L)[c]	13 ± 29		14 ± 36		12 ± 19		0.0164
White blood cell count (n/µl)	7347 ± 3450		7499 ± 3352		7178 ± 3552		0.2452
Serum creatinine at index biopsy (µmol/L)	210 (162, 301)		204 (168, 286)		214 (156, 318)		0.7728
eGFR at index biopsy (ml/min)	37 ± 18		38 ± 17		35 ± 19		0.0458
Baseline serum creatinine (µmol/L)	178 (133, 251)		182 (135, 239)		172 (128, 273)		0.6012
Hydronephrosis, any grade (unknown)	20 (210)	3.2 (33.1)	15 (45)	4.6 (13.7)	5 (165)	1.6 (53.9)	0.5636
Transplant renal arterial stenosis	18	2.8	1	0.3	17	5.6	<0.0001

[a]91 missing values.
[b]Overall *P* value vs none.
[c]24 missing values.
"Baseline serum creatinine" refers to the lowest value within 30 d before the index biopsy.
eGFR, estimated glomerular filtration rate.

and cases without TCMR is illustrated in Figure 4A. Using the prespecified MS classifier cutoff of –0.25, tubulointerstitial TCMR was recognized in 67% and vascular TCMR in 65% as rejection by the urine test (Figure 4B). Borderline TCMR was classified as rejection in a similar frequency as cases without TCMR. Additional glomerulitis and peritubular capillaritis only numerically increased positive results of the classifier in TCMR cases grade I–III ($P = 0.16$). However, including cases with other rejection findings than TCMR I–III into the rejection group, namely glomerulitis and peritubular capillaritis or borderline TCMR, decreased the test performance (Table 5). Similarly, excluding borderline TCMR from the analyses did not improve the test performance. Further analyses in subgroups (Table 5) showed that the performance of the urine test was higher in samples taken within the first 6 wk of posttransplantation and in female subjects. There were transplant center-related differences, with highest test performance with samples from Leuven and Hannover. Lower performance was observed for cases with minimal adequate biopsies.

The MS classifier score correlated weakly with the serum creatinine at biopsy within the whole group of patients (r = 0.28 [95% CI, 0.20-0.35; $P < 0.0001$]), in patients without TCMR (r = 0.30 [95% CI, 0.20-0.39; $P < 0.0001$]), but not in patients with TCMR (r = 0.20 [95% CI, 0.03-0.40; $P = 0.07$]). The MS classifier did not correlate with the percentage increase in serum creatinine at biopsy compared with baseline values. Other associations with the MS classifier in univariable analyses indicated potential confounding effects. Generally, higher classifier values were observed in patients with infections including urinary tract infection and with heparin and calcium supplement treatments, lower values with statin treatment. Higher classifier values in patients without rejection were observed in association with severe acute tubular injury, delayed graft function, in male recipients and recipients of female donor organs. Yet, sensitivity analyses with these variables did not indicate relevant effects on the performance of the urine test (**Table S4, SDC,** http://links.lww.com/TXD/A411). Differences in urine volume and concentration might have affected peptide marker amplitudes. However, reanalysis considering urinary creatinine concentration did not change the classification performance of the marker set (**Table S5, SDC,** http://links.lww.com/TXD/A411).

According to the protocol of this in-place validation study, the reference standard for comparison with the index test was the biopsy result reported by the local pathologist. To assess whether heterogeneity of this evaluation contributed to the low performance of the index test, 409 biopsies were secondarily re-evaluated by 2 nephropathologists (J.H.B. and A.K.) to obtain an agreed diagnosis. Interobserver agreement between the primary and secondary biopsy evaluation was low, with a Krippendorff's alpha of 0.38 (95% CI, 0.27-0.48) over all diagnosis categories (Figure 5) and 0.49 (95% CI, 0.32-0.62) for TCMR grade I–III versus no TCMR or borderline TCMR.

TABLE 3.

Timing of biopsies and histomorphological results

	Total samples n = 635	%	Prospectively recruited n = 329	%	Archived samples n = 306	%	*P*
Biopsies during weeks 1 and 2 after Tx	203	32.0	104	31.6	99	32.4	0.1967
Biopsies during weeks 3 and 4 after Tx	89	14.0	43	13.1	46	15.0	
Biopsies during weeks 5 and 6 after Tx	46	7.2	18	5.5	28	9.2	
Biopsies after week 6	296	46.7	163	49.5	133	43.5	
Fully adequate biopsies	371	58.4	119	36.2	252	82.4	<0.0001
Minimal adequate biopsies	119	18.7	82	24.9	37	12.1	
Inadequate biopsies	143	22.5	126	38.3	17	5.6	
Unknown biopsy adequacy	2	0.3	2	0.6	0	—	
Acute TCMR							0.0503
None	390	62.0	198	61.3	192	62.7	
Borderline	157	24.7	87	26.4	70	22.9	
TCMR	82	12.9	38	11.6	44	14.4	
IA	21	3.3	11	3.3	10	3.2	
IB	15	2.4	7	2.1	8	2.6	
IIA	37	5.8	12	3.7	25	8.2	
IIB	8	1.3	7	2.1	1	0.3	
III	1	0.2	1	0.3	0	0	
Time of acute TCMR including borderline cases							0.3938
Weeks 1 and 2 after Tx	103	16.2	52	15.8	51	16.6	
Weeks 3 and 4 after Tx	35	5.5	15	4.6	19	6.2	
Weeks 5 and 6 after Tx	11	1.7	5	1.5	6	2.0	
After week 6	96	15.1	55	16.7	38	12.4	
Time of acute TCMR excluding borderline cases							0.7755
Weeks 1 and 2 after Tx	38	6.0	19	5.8	19	6.2	
Weeks 3 and 4 after Tx	10	1.6	4	1.2	6	2.0	
Weeks 5 and 6 after Tx	3	0.5	2	0.6	1	0.3	
After week 6	31	4.9	13	3.4	18	5.9	
Acute antibody-mediated rejection features							
Glomerulitis	104	16.4	27	8.2	77	25.2	<0.0001
Peritubular capillaritis	80	12.6	29	8.8	51	16.7	0.0030
C4d positivity[a]	93	14.6	26	7.9	67	21.9	<0.0001
C4d focal	55		12		43		
C4d diffuse	36		12		24		
Thrombotic microangiopathy	7	1.1	6[b]	1.8	1[b]	0.3	0.0711
Transplant glomerulopathy	7	1.1	4	1.2	3	1.0	1.0000
Transplant vasculopathy	17	2.7	5	1.5	12	3.9	0.0836
Total i-score							<0.0001
0	182	65.2	76	48.7	106	86.2	
1	67	24.0	56	35.9	11	8.9	
2	18	6.5	12	7.7	6	4.9	
3	12	4.3	12	7.7	0	0	
Unknown	356		173		183		
Interstitial fibrosis and tubular atrophy							<0.0001
Grade 0	310	50.5	125	40.2	185	61.1	
Grade I	234	38.1	163	52.4	71	23.4	
Grade II	47	7.7	16	5.1	31	10.2	
Grade III	23	3.7	7	2.3	16	5.3	
Unknown	21		18		3		
Acute tubular injury							<0.0001
None	147	23.7	45	14.3	102	33.3	
Mild/focal	258	41.6	170	54.1	88	28.8	
Moderate-severe/diffuse	215	34.7	99	31.5	116	37.9	
Unknown	15		15		0		
Isometric tubular vacuolization	25	3.9	8	2.4	17	5.5	0.0641
BK virus nephropathy	26	4.2	18	5.7	8	2.6	<0.0001
Glomerulonephritis	12	1.9	10	3.0	2	0.7	0.0383
Nephrosclerosis	98	15.4	84	25.5	14	4.5	<0.0001
Ascending infection	8	1.3	6	1.8	2	0.7	0.1865

[a]C4d positivity in peritubular capillaries in 2 cases without exact grading, 12 cases of C4d positivity with ABO-incompatible transplantation.
[b]One case each with additional glomerulitis and peritubular capillaritis. For variables with multiple categories, *P* denotes the results of a chi-square test over all categories between prospectively vs archived samples.
Adequacy was determined according to the criteria of the Banff classification.
TCMR, T cell–mediated rejection; Tx, transplantation.

TABLE 4.

Inventory of TCMR cases

	All cases	With additional glomerulitis		With additional peritubular capillaritis		With additional glomerulitis and peritubular capillaritis	
	N	n	%	n	%	N	%
No TCMR	390	29	7.4	9	2.3	27	6.9
Borderline TCMR	157	12	7.6	5	3.2	12	7.6
TCMR IA, IB	36	1	2.8	8	22.2	7	19.4
TCMR IIA, IIB, III	46	8	17.4	4	8.7	8	17.4

"All cases" denotes the total number of biopsies with the different acute TCMR phenotypes and without TCMR, percentages are row percentages.
TCMR, T cell–mediated rejection.

Performance of the index test improved slightly with this secondary assessment, with an AUC of 0.63 instead of 0.60, a sensitivity of 0.73 instead of 0.66 (95% CI, 0.58-0.88), and a specificity of 0.45 instead of 0.47 (95% CI, 0.40-0.50).

According to the study protocol, data from a follow-up observation were included in a secondary analysis. The rationale was that an acute rejection episode might have been missed by the index biopsy eg, because of sampling error, and thus left untreated but then was detected by a short-term follow-up biopsy. Forty-seven patients without TCMR in the index biopsy had an acute rejection in the following 2 mo. Inclusion of these cases into the patient group with acute rejection in the index biopsy did not improve the diagnostic performance of the MS marker set (AUC 0.53, sensitivity 0.58 [95% CI, 0.50-0.67], specificity 0.46 [95% CI, 0.42-0.51]).

In a post hoc analysis, the behavior of single peptides of the marker set was compared between the training cohort used to establish the marker set and the validation cohort reported here (**Table S6, SDC,** http://links.lww.com/TXD/A411). Five downregulated peptides of the training set were also significantly downregulated in the validation cohort and 2 peptides were upregulated in both cohorts. Six peptides showed no significant upregulation or downregulation in the validation cohort and 1 peptide had a discordant behavior between the 2 cohorts. Application of a marker set that contains only the 7 peptides with concordant behavior (with unchanged support vector machine settings) increased the AUC significantly from 0.60 to 0.67 (P = 0.009) in the validation cohort (**Figure S1, SDC,** http://links.lww.com/TXD/A411). Replicability of the peptide marker set was also tested on an independent cohort of 690 patients from another international study[19] (BIOMARGIN; ClinicalTrials.gov number NCT02832661), showing an AUC of 0.63 with the original marker set and of 0.69 with the reduced marker set, which was again a significant improvement (P = 0.005; **Figure S2, SDC,** http://links.lww.com/TXD/A411). Five peptides of the marker set in this additional validation cohort were consistently downregulated and 2 peptides upregulated, as compared with the training cohort (**Table S6, SDC,** http://links.lww.com/TXD/A411). **Table S6** also gives the parent proteins of single peptides that were identifiable by sequencing, revealing different collagens as origin.

DISCUSSION

This multicenter, in-place diagnostic study aimed to validate a predefined urinary peptide marker set to detect acute TCMR in kidney transplant recipients. The marker set failed to achieve the expected performance that is required for practical application.

Several limitations need to be addressed. First, the predefined sample size of at least 600 patients was not achieved in the planned study period. Lower recruitment may be in part explained by the fact that several centers had changed to an immunosuppressive protocol with tacrolimus at study begin, leading to a lower biopsy rate due to less rejections. Lacking recruitment was compensated by including archived samples that had been collected according to the same protocol as for the prospectively recruited patients. Patients with prospectively collected and archived samples differed in several aspects, including proportions of living donor and AB0 blood group-incompatible transplantations and differently intense induction therapies. This may in part, reflect center-specific practices, but also differences in patient's immunological risk profile among the centers. In fact, archived samples, which were mainly derived from Leuven, Paris, and Hannover, showed more vascular TCMRs. Potential bias introduced by inclusion of archived samples was countered by separate description of clinical and laboratory variables and a sensitivity analysis regarding the performance measures of the urine test. Despite sufficient overall numbers of patients for the analysis, the precalculated number of TCMRs was lower than expected, with only 82 instead of 150 cases. Decreasing incidence of TCMR has been reported and appears to be related to the increasing use of tacrolimus in combination with mycophenolate mofetil.[20-22] The low number of TCMR may have lowered the sensitivity of the study and limited in-depth analysis of subgroups with different rejection severity. Finally, it was planned to re-evaluate all biopsies centrally, but organizational reasons limited this to two-thirds of the biopsies.

The performance of the urine test is certainly too low to predict TCMR reliably or to support clinicians in deciding whether to perform a biopsy in patients with graft dysfunction. Based on the observed sensitivity, specificity and incidence of acute TCMR in the whole study cohort, 34% of the rejection cases would have been missed and in 53% of cases without rejection, the test would have suggested performing a biopsy to confirm the presence of rejection.

Separate analysis of tubulointerstitial and vascular TCMR cases revealed similar detection rates of 67% and 65%. Borderline rejections were classified similarly frequent as rejection as the control samples without TCMR. The meaning of borderline rejection has been debated in recent years, leading to changes in the thresholds of histomorphological criteria by the Banff group.[23,24] Molecular studies suggested that borderline findings rather represent nonspecific injury than true rejection, especially in early biopsies.[25] Conversely, cellular infiltrates in the tubulointerstitial compartment, even when presenting below the defined thresholds for establishing the diagnosis of TCMR or borderline rejection, have

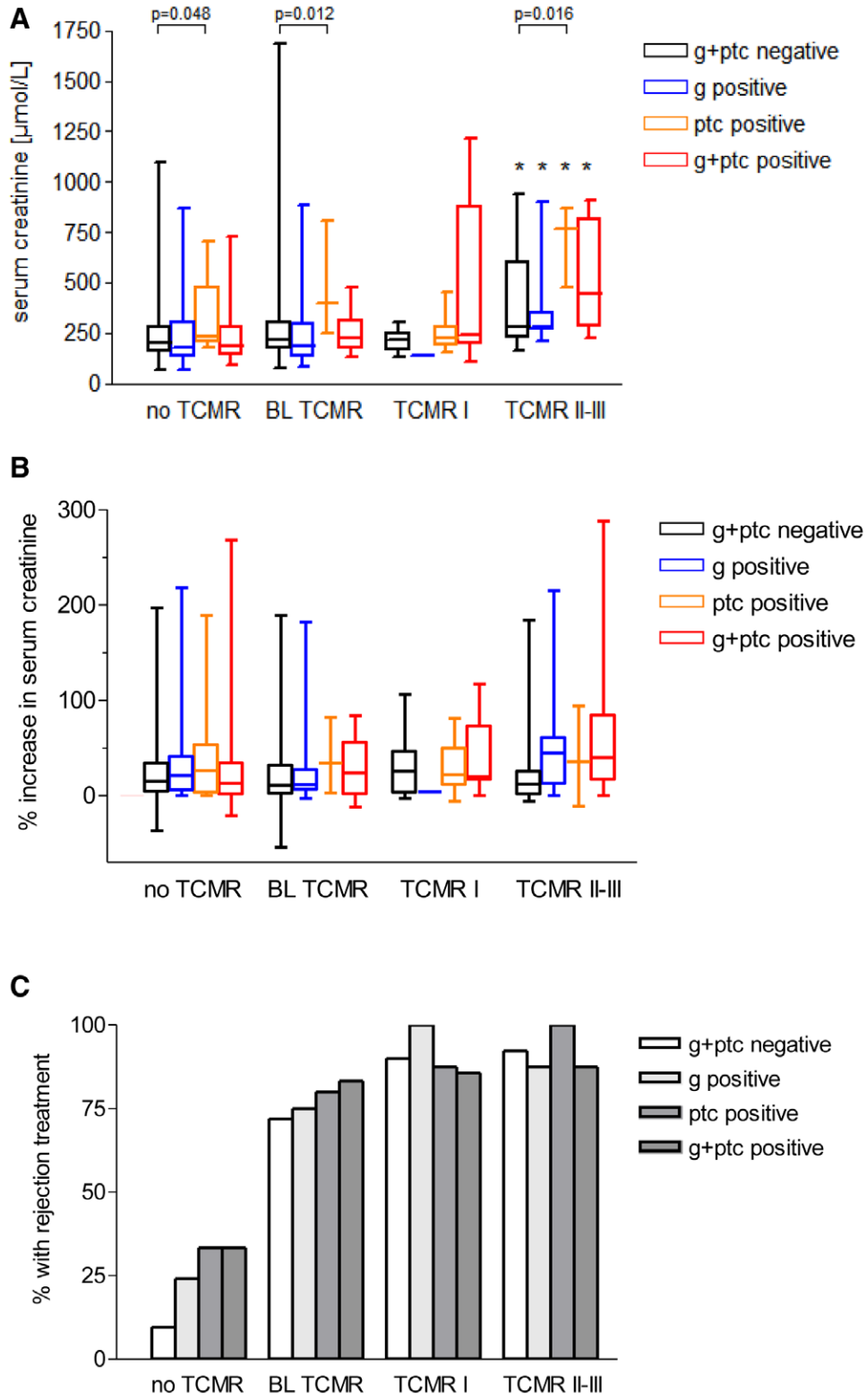
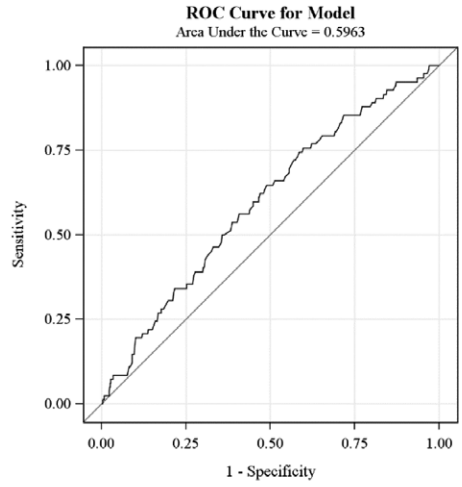


FIGURE 2. Graft function and rejection treatment in different TCMR grades. A, Serum creatinine concentration at the time of biopsy (*$P < 0.0001$ in g+ptc negative, $P = 0.031$ in g positive, $P = 0.011$ in ptc positive, $P = 0.038$ in g+ptc positive TCMR grade II–III cases vs the corresponding groups without TCMR). B, Percentage increase in serum creatinine at the time of biopsy compared with the lowest value within the 30 d before biopsy. C, Proportion of cases with rejection treatments. Boxes and whiskers represent medians, lower and upper quartiles, and the extreme values. BL, borderline; g, glomerulitis of any grade; ptc, peritubular capillaritis of any grade; TCMR, T cell–mediated rejection.

been associated with an inferior long-term graft outcome.[26,27] Coexisting glomerulitis and/or peritubular capillaritis in cases with TCMR tended to increase the rate of samples classified positive for rejection. This could be because of an overall more severe rejection process and inflammation, which was recognized by the peptide marker set more easily. Based on



	AUC	Sensitivity (95%-CI)	Specificity (95%-CI)	PPV (95%-CI)	NPV (95%-CI)
Total	0.60	0.66 (0.56-0.76)	0.47 (0.43-0.51)	0.16 (0.12-0.20)	0.90 (0.87-0.94)
Prospective	0.57	0.63 (0.47-0.78)	0.47 (0.41-0.52)		
Archived	0.62	0.68 (0.54-0.82)	0.47 (0.41-0.53)		

FIGURE 3. Performance of the urinary peptide test to diagnose acute TCMR. The ROC AUC is shown for the entire patient cohort with 82 cases of TCMR and separately for prospectively recruited patients and patients with archived samples. AUC, area under the curve; CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value; TCMR, T cell–mediated rejection.

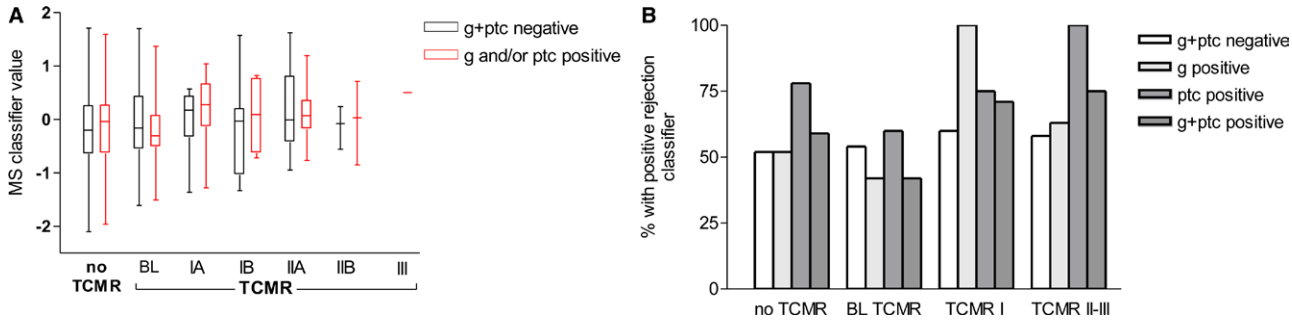


FIGURE 4. MS classifier results in patients with and without TCMR. A, Urinary peptide classifier scores of different TCMR grades compared with no TCMR. B, Percentage of rejection-positive classifier results in cases with and without TCMR. Red bars denote samples with additional glomerulitis (g) and/or peritubular capillaritis (ptc), black bars cases without. Boxes and whiskers represent medians, lower and upper quartiles, and the extreme values. BL, borderline; MS, mass spectrometry; TCMR, T cell–mediated rejection.

the observed high prevalence of glomerulitis and peritubular capillaritis in TCMR cases, marker sets for acute rejection are perhaps clinically most useful when they are able to detect characteristics of both, TCMR and AMR. The high prevalence of combined TCMR/AMR within the first transplant year was also noted in a recent large registry study.[2]

The low performance of the urine test suggests that specific features of TMCR were not detected with sufficient sensitivity and, on the other hand, that nonspecific injury was recognized by the test. Because of the nature of this study, with biopsy at the time of graft damage ("biopsies for cause"), a high proportion of cases without TCMR had general injury features not specific for rejection, like acute tubular injury in 76% and interstitial inflammation in 35%. This was also reflected by similar impairment of graft function at biopsy in patients with and without rejection. In univariate analyses, the presence of urinary tract infection and systemic infection (indicated by leukocytosis, elevated C-reactive protein, and antibiotic treatment) increased the MS classifier value toward the rejection diagnosis. Also, severe acute tubular injury and delayed graft function due to any cause were weakly associated with a higher MS classifier value. Yet, sensitivity analyses accounting for these conditions in subgroups did not indicate clinically relevant effects on the overall test performance. When developing the urine test, care was taken to have these unspecific injuries sufficiently represented in the training set, particularly in the control samples without rejection. Nonetheless, the peptide marker set contained mostly collagen fragments (**Table S6, SDC,** http://links.lww.com/TXD/A411) similar to another study,[28] which could be an indication of any type of damage to the nephron, thus rendering the urine test too unspecific to separate such injury from rejection. Despite these difficulties, the post hoc analysis of individual markers with optimization of the marker set shows that improved

TABLE 5.

Sensitivity analysis in subgroups

	AUC	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Primary: No TCMR (incl. borderline) vs TCMR I–III	0.60	0.66 (0.56-0.76)	0.47 (0.43-0.51)	0.16 (0.12-0.20)	0.90 (0.87-0.94)
No TCMR vs TCMR I–III[a]	0.61	0.65 (0.55-0.76)	0.48 (0.42-0.53)	0.34 (0.18-0.30)	0.85 (0.79-0.90)
No TCMR vs borderline TCMR and TCMR I–III	0.56	0.57 (62.1-0.64)	0.48 (0.42-0.53)	0.45 (0.39-0.50)	0.60 (0.54-0.66)
No rejection vs TCMR I–III and cases with AMR signs[a]	0.57	0.62 (0.55-0.70)	0.48 (0.42-0.53)	0.35 (0.30-0.29)	0.74 (0.68-0.80)
No rejection vs borderline TCMR, TCMR I–III and cases with AMR signs	0.55	0.58 (0.52-0.63)	0.48 (0.42-0.53)	0.51 (0.45-0.56)	0.55 (0.49-0.60)
Male recipients	0.55	0.59 (0.46-0.72)	0.44 (0.38-0.51)	0.22 (0.15-0.28)	0.81 (0.74-0.88)
Female recipients	0.72	0.78 (0.62-0.94)	0.54 (0.44-0.63)	0.28 (0.18-0.39)	0.91 (0.84-0.98)
Leuven	0.63	0.72 (0.56-0.89)	0.43 (0.28-0.57)	0.44 (0.30-0.58)	0.71 (0.55-0.88)
Paris	0.55	0.55 (0.25-0.84)	0.49 (0.36-0.62)	0.17 (0.05-0.30)	0.85 (0.73-0.97)
Hannover	0.72	0.86 (0.71-1.00)	0.48 (0.39-0.56)	0.22 (0.13-0.31)	0.95 (0.90-1.00)
Other German centers besides Hannover	0.50	0.42 (0.20-0.64)	0.53 (0.42-0.65)	0.19 (0.07-0.31)	0.78 (0.67-0.90)
Biopsies within the first 6 wk post-Tx	0.63	0.73 (0.60-0.85)	0.41 (0.33-0.49)	0.29 (0.21-0.37)	0.82 (0.73-0.90)
Biopsies after 6 wk post-Tx	0.56	0.53 (0.36-0.71)	0.54 (0.46-0.61)	0.17 (0.09-0.25)	0.87 (0.80-0.93)
Fully adequate biopsies	0.62	0.67 (0.54-0.81)	0.45 (0.38-0.53)	0.24 (0.17-0.31)	0.85 (0.77-0.92)
Minimal adequate biopsies	0.55	0.50 (0.24-0.76)	0.48 (0.36-0.60)	0.17 (0.06-0.29)	0.82 (0.69-0.94)
Not adequate biopsies	0.63	0.71 (0.52-0.91)	0.53 (0.42-0.65)	0.31 (0.18-0.44)	0.87 (0.77-0.97)

[a]With exclusion of borderline TCMR from the analysis. Subanalyses were performed using the primary rejection diagnosis (no TCMR including borderline TCMR vs TCMR grade I–III) unless otherwise stated.
Values in parentheses denote the 95% confidence intervals.
AMR, antibody-mediated rejection; TCMR, T cell–mediated rejection.

performance can generally be achieved, even if this was still not quite sufficient for the requirements of clinical use (**Table S6, Figures S1 and S2, SDC,** http://links.lww.com/TXD/A411).

The performance of any experimental, alternative test that is evaluated against the established reference standard is directly dependent on the reliability of that standard. The results of the independent re-evaluation of >400 biopsies by 2 pathologists agreeing on a common diagnosis confirmed the moderate interobserver agreement noted in earlier studies of kidney graft biopsies.[29] However, the probably more homogeneous reassessment of biopsies only led to a slight increase in the AUC and sensitivity of the urine test. Nonrepresentative biopsies might also have contributed to an unreliable histomorphological diagnosis by missing the rejection in too small samples. Detailed analysis of fully, minimal and not adequate biopsies showed no systematic trend toward decreasing sensitivity and specificity of the urine test, indicating that biopsy adequacy was not an important factor for the poor test performance. Also, missed rejection diagnosis in the evaluation of the index biopsy appeared to be not relevant as indicated by the inclusion of rejection episodes of short-term follow-up biopsies in the analysis.

Recent reviews have summarized the studies that employed urine protein and peptide markers for the detection of rejections.[3-5] Numerous studies established and explored combinations of a few peptides/proteins such as granzyme B, CXCL-9, CXCL-10,[30-34] or more complex proteomic marker sets.[3,4] In approximately a third, reliability of the markers was examined independently on separate samples. Yet, since these were basically selected samples, this represents no real in-place validation.[3] Thus, despite the array of putative rejection markers, there is, to our knowledge, currently no truly validated test system with proteomic markers in urine that is in widespread clinical use or commercially available. At the beginning of this study, another 4 studies with comparable proteomic approaches were listed at ClinicalTrials.gov. One (NCT01515605) began in 2011 with the goal to recruit 1000 kidney transplant patients until 2014 for examination of proteomic markers and specific molecules in blood and urine in a longitudinal fashion but is still ongoing. Another

	Re-assessment of the biopsies						
First assessment	No rejection	Borderline TCMR	TCMR IA	TCMR IB	TCMR IIA	TCMR IIB	TCMR III
No rejection	230	9	2	0	3	2	0
Borderline TCMR	74	34	1	0	1	0	2
TCMR IA	3	2	5	1	1	1	0
TCMR IB	3	0	2	3	0	0	0
TCMR IIA	15	4	0	1	5	1	0
TCMR IIB	0	0	0	0	1	1	0

FIGURE 5. Interobserver agreement on rejection diagnosis. Number of cases are depicted. Shaded fields indicate agreement. First assessment: diagnosis from the local pathologist from each center; reassessment: diagnosis from 2 pathologists, with agreement on the diagnosis after masked evaluation. TCMR, T cell–mediated rejection.

study (NCT01289717) recruited 307 kidney transplant recipients from 2011 until 2016 to evaluate proteomic and other molecular markers in blood, urine, and graft biopsies for early detection of rejection. The results of this study are pending. A small study (NCT02463253) with 20 kidney transplant recipients was begun in 2015 to study proteogenomic and proteomic biomarkers in blood, urine, and biopsies with acute rejection and chronic lesions. A large study (NCT 01531257) began in 2010 to recruit 1000 kidney transplant recipients, aiming at validation of proteogenomic biomarkers for acute rejection and chronic lesions in blood, urine, and graft biopsies. Study completion is expected in the year 2025. Another study (NCT02832661; BIOMARGIN) established proteomic biomarkers for AMR that were highly accurate in an independent, unselected validation cohort but failed to establish proteomic biomarkers for TCMR of sufficient diagnostic performance.[35] Using a multiparametric model that included results on 2 urinary chemokines, CXCL-9 and CXCL-10, AMR and TCMR were detected with sufficient diagnostic performance.[33] Based on this scarcity of proteomic data with proven sufficient, sensitivity and specificity to detect TCMR future results must be awaited.

In view of the negative result of this study, some learning points can be derived. General considerations concern sufficiently large validation cohorts with realistic numbers of recruited index cases, which represent the whole spectrum and heterogeneity of rejection. Robustness of the gold standard is another point that needs consideration, eg, by planning consent reading of biopsies by pathology experts. Regarding marker development, the high prevalence of biopsies with histomorphological characteristics of AMR suggests that marker sets that can detect TCMR as well as AMR and mixed rejection cases are most advantageous. Training sets of sufficient size should reflect the whole spectrum of nonspecific injuries and other confounders of the test in controls and cases. As illustrated in this study, a stepwise approach with testing and optimizing markers in 1 validation cohort and application to the next validation cohort can improve diagnostic performance.

ACKNOWLEDGMENTS

The advisory support during the study by Prof. Christian Hugo, Department of Nephrology, University Dresden, Germany and by Prof. Hans-Heinrich Kreipe, Department of Pathology, Germany, is acknowledged.

REFERENCES

1. El Ters M, Grande JP, Keddis MT, et al. Kidney allograft survival after acute rejection, the value of follow-up biopsies. *Am J Transplant.* 2013;13:2334–2341.
2. Clayton PA, McDonald SP, Russ GR, et al. Long-term outcomes after acute rejection in kidney transplant recipients: an ANZDATA analysis. *J Am Soc Nephrol.* 2019;30:1697–1707.
3. Gwinner W, Metzger J, Husi H, et al. Proteomics for rejection diagnosis in renal transplant patients: where are we now? *World J Transplant.* 2016;6:28–41.
4. Anglicheau D, Naesens M, Essig M, et al. Establishing biomarkers in transplant medicine: a critical review of current approaches. *Transplantation.* 2016;100:2024–2038.
5. Menon MC, Murphy B, Heeger PS. Moving biomarkers toward clinical implementation in kidney transplantation. *J Am Soc Nephrol.* 2017;28:735–747.
6. Metzger J, Chatzikyrkou C, Broecker V, et al. Diagnosis of subclinical and clinical acute T-cell-mediated rejection in renal transplant patients by urinary proteome analysis. *Proteomics Clin Appl.* 2011;5:322–333.
7. Zapf A, Gwinner W, Karch A, et al. Non-invasive diagnosis of acute rejection in renal transplant patients using mass spectrometry of urine samples - a multicentre phase 3 diagnostic accuracy study. *BMC Nephrol.* 2015;16:153.
8. Steering Committee of the Istanbul Summit. Organ trafficking and transplant tourism and commercialism: the Declaration of Istanbul. *Lancet*. 2008;372:5–6.
9. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310:2191–2194.
10. Jantos-Siwy J, Schiffer E, Brand K, et al. Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease. *J Proteome Res.* 2009;8:268–281.
11. Mischak H, Vlahou A, Ioannidis JP. Technical aspects and inter-laboratory variability in native peptide profiling: the CE-MS experience. *Clin Biochem.* 2013;46:432–443.
12. Siwy J, Mullen W, Golovko I, et al. Human urinary peptide database for multiple disease biomarker discovery. *Proteomics Clin Appl.* 2011;5:367–374.
13. Klein J, Papadopoulos T, Mischak H, et al. Comparison of CE-MS/MS and LC-MS/MS sequencing demonstrates significant complementarity in natural peptide identification in human urine. *Electrophoresis.* 2014;35:1060–1064.
14. Pejchinovski M, Klein J, Ramírez-Torres A, et al. Comparison of higher energy collisional dissociation and collision-induced dissociation MS/MS sequencing methods for identification of naturally occurring peptides in human urine. *Proteomics Clin Appl.* 2015;9:531–542.
15. Zürbig P, Renfrow MB, Schiffer E, et al. Biomarker discovery by CE-MS enables sequence analysis via MS/MS with platform-independent separation. *Electrophoresis.* 2006;27:2111–2125.
16. Mengel M, Sis B, Haas M, et al; Banff Meeting Report Writing Committee. Banff 2011 Meeting report: new concepts in antibody-mediated rejection. *Am J Transplant.* 2012;12:563–570.
17. Haas M, Sis B, Racusen LC, et al; Banff Meeting Report Writing Committee. Banff 2013 meeting report: inclusion of c4d-negative antibody-mediated rejection and antibody-associated arterial lesions. *Am J Transplant.* 2014;14:272–283.
18. Tawab KA, Gheith O, Al Otaibi T, et al. Recurrent urinary tract infection among renal transplant recipients: risk factors and long-term outcome. *Exp Clin Transplant.* 2017;15:157–163.
19. Van Loon E, Gazut S, Yazdani S, et al. Development and validation of a peripheral blood mRNA assay for the assessment of antibody-mediated kidney allograft rejection: a multicentre, prospective study. *EbioMedicine.* 2019;46:463–472.
20. Michielsen LA, van Zuilen AD, Verhaar MC, et al. Effect of initial immunosuppression on long-term kidney transplant outcome in immunological low-risk patients. *Nephrol Dial Transplant.* 2019;34:1417–1422.
21. Wagner M, Earley AK, Webster AC, et al. Mycophenolic acid versus azathioprine as primary immunosuppression for kidney transplant recipients. *Cochrane Database Syst Rev*. 2015:1–122.
22. Jones-Hughes T, Snowsill T, Haasova M, et al. Immunosuppressive therapy for kidney transplantation in adults: a systematic review and economic model. *Health Technol Assess.* 2016;20:1–594.
23. Loupy A, Haas M, Roufosse C, et al. The Banff 2019 Kidney Meeting Report (I): updates on and clarification of criteria for T cell– and antibody-mediated rejection. *Am J Transplant.* 2020;20:2318–2331.
24. 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.
25. Mengel M, Chang J, Kayser D, et al. The molecular phenotype of 6-week protocol biopsies from human renal allografts: reflections of prior injury but not future course. *Am J Transplant.* 2011;11:708–718.
26. Mengel M, Gwinner W, Schwarz A, et al. Infiltrates in protocol biopsies from renal allografts. *Am J Transplant.* 2007;7:356–365.
27. Mehta RB, Tandukar S, Jorgensen D, et al. Early subclinical tubulitis and interstitial inflammation in kidney transplantation have adverse clinical implications. *Kidney Int.* 2020;98:436–447.

28. Ling XB, Sigdel TK, Lau K, et al. Integrative urinary peptidomics in renal transplantation identifies biomarkers for acute rejection. *J Am Soc Nephrol.* 2010;21:646–653.
29. Marcussen N, Olsen TS, Benediktsson H, et al. Reproducibility of the Banff classification of renal allograft pathology. Inter- and intraobserver variation. *Transplantation.* 1995;60:1083–1089.
30. Li B, Hartono C, Ding R, et al. Noninvasive diagnosis of renal-allograft rejection by measurement of messenger RNA for perforin and granzyme B in urine. *N Engl J Med.* 2001;344:947–954.
31. Matz M, Beyer J, Wunsch D, et al. Early post-transplant urinary IP-10 expression after kidney transplantation is predictive of short- and long-term graft function. *Kidney Int.* 2006;69:1683–1690.
32. Tatapudi RR, Muthukumar T, Dadhania D, et al. Noninvasive detection of renal allograft inflammation by measurements of mRNA for IP-10 and CXCR3 in urine. *Kidney Int.* 2004;65:2390–2397.
33. Tinel C, Devresse A, Vermorel A, et al. Development and validation of an optimized integrative model using urinary chemokines for noninvasive diagnosis of acute allograft rejection. *Am J Transplant.* 2020;20:3462–3476.
34. Hu H, Aizenstein BD, Puchalski A, et al. Elevation of CXCR3-binding chemokines in urine indicates acute renal-allograft dysfunction. *Am J Transplant.* 2004;4:432–437.
35. Mertens I, Willems H, Van Loon E, et al. Urinary protein biomarker panel for the diagnosis of antibody-mediated rejection in kidney transplant recipients. *Kidney Int Rep.* 2020;5:1448–1458.