REVIEW

The 1-year Renal Biopsy Index: a scoring system to drive biopsy indication at 1-year post-kidney transplantation

Magali Giral[1,2], Karine Renaudin[3], Maarten Naesens[4,5], Redmer Luning[4,5], Dany Anglicheau[6], Emmanuel Morelon[7], Alexandre Huneau[8], Chloé Paul[1,2], Sophie Brouard[1,2], Grégoire Couvrat-Desvergnes[1,9], Yohann Foucher[8,10] & Etienne Dantan[8]

1 CRTI UMR 1064, Inserm, Université de Nantes, ITUN, CHU Nantes, RTRS Centaure, Nantes, France
2 Centre d'Investigation Clinique en Biothérapie, Labex Transplantex, Nantes, France
3 Pathological Anatomy and Cytology, CHU Hôtel-Dieu, Nantes, France
4 Department of Microbiology and Immunology, KU Leuven, Leuven, Belgium
5 Department of Nephrology and Renal Transplantation, University Hospitals Leuven, Leuven, Belgium
6 Kidney Transplant Center, Necker University Hospital, APHP, RTRS "Centaure", Paris Descartes and Sorbonne Paris Cité Universities, Paris, France
7 Nephrology, Transplantation and Clinic Immunology Department, RTRS "Centaure", Edouard Herriot University Hospital, Hospices Civils, Lyon, France
8 INSERM UMR 1246 – SPHERE, Nantes University, Tours University, Nantes, France
9 Department of Nephrology, Dialysis and Transplantation, Departmental Hospital of Vendée, La Roche-sur-Yon, France
10 CHU NANTES, Nantes, France

SUMMARY

Surveillance biopsies after renal transplantation remain debatable. To drive the decision of such intervention, we propose a predictive score of abnormal histology at 1-year post-transplantation, named 1-year Renal Biopsy Index (1-RBI). We studied 466 kidney recipients from the DIVAT cohort alive with a functioning graft and a surveillance biopsy at 1-year post-transplantation. Patients displaying abnormal histology (49%) (borderline, acute rejection, interstitial fibrosis and tubular atrophy [IFTA] grade 2 or 3, glomerulonephritis) were compared to the normal or subnormal (IFTA grade 1) histology group. Obtained from a lasso penalized logistic regression, the 1-RBI was composed of recipient gender, serum creatinine at 3, 6, and 12 month post-transplantation and anticlass II immunization at transplantation (internal validation: AUC = 0.71, 95% CI [0.53–0.83]; external validation: AUC = 0.62, 95% CI [0.58–0.66]). While we could not determinate a threshold able to identify patients at high chance of normal or subnormal histology, we estimated and validated a discriminating threshold capable of identifying a subgroup of 15% of the patients with a risk of abnormal histology higher than 80%. The 1-RBI is computable online at www.divat.fr. The 1-RBI could be a useful tool to standardize 1-year biopsy proposal and may for instance help to indicate one in case of high risk of abnormal histology.

Transplant International 2018; 31: 947–955

Key words
histological diagnostic, kidney transplantation, predictive score, surveillance biopsy

Received: 19 February 2018; Revision requested: 26 March 2018; Accepted: 7 June 2018

Correspondence
Magali Giral, Centre de Recherche en Transplantation et Immunologie INSERM UMR1064, Université de Nantes, Centre Hospitalier Universitaire de Nantes, 30 Bd Jean Monnet, Nantes, France.
Tel.: 33 2 40 08 74 43;
fax: 33 2 40 08 74 11;
e-mail: mgiral@chu-nantes.fr

doi:10.1111/tri.13290

Introduction

Surveillance biopsies within the first year post-kidney transplantation are increasingly performed as they allow to identify subclinical graft lesions, which are associated with long-term outcomes [1–4]. Surveillance biopsies provide precious histological information on the mechanism and physio-pathologic knowledge.

Nevertheless, one can list several arguments in favor of the choice made by numerous physicians not to propose surveillance biopsies. Surveillance biopsies are costly and invasive with possible serious adverse events [5]. It also remains controversial whether 1-year surveillance biopsies provide long-term clinical benefits as studies generally used short-term endpoints [6–8], such as modification in immunosuppressive drugs or renal function changes after 6 months following the biopsy [9]. There is actually no therapeutic consensus following the identification of most lesions such as inflammatory fibrosis or isolated microcirculation inflammation without donor-specific antibodies (DSA) [10,11]. There is an important percentage of normal histological results on 1-year surveillance biopsies with no indication of therapeutic changes and for which it has been shown that histological deterioration is rarely observed afterward [12,13]. Finally, the choice of the optimal timing to perform the biopsy and the necessity to repeat it is not clearly recommended. Early biopsies at 1 or 3 months generally focus on subclinical rejection or ischemia reperfusion injuries [14], in contrary to later biopsies at 6 months or 1 year that mainly screen for interstitial fibrosis and tubular atrophy (IFTA) lesions progression or calcineurin inhibitor (CNI) toxicity [15].

As a consequence, it appears essential to reduce the observed heterogeneity of practices: on one hand, the physicians convinced that surveillance biopsies constitute a useful tool of graft monitoring, on the other hand, those who believed that the adverse events counterbalance the unproven utility of surveillance biopsies. In this context, we hypothesized that a noninvasive and clinical-based diagnostic test of graft lesions could help physicians to recommend biopsies more specifically in patients more susceptible to be treated. The objective of our study was to define and validate such a diagnostic test at 1-year post-transplantation.

Patients and methods

Studied cohorts

We considered adult recipients, transplanted from the first or second kidney of a heart-beating deceased donor, alive with a functioning graft at 1-year post-transplantation and with a 1-year surveillance biopsy performed with a complete Banff classification, without other transplanted organs. We voluntarily not included patients with combined organs since these patients differed from the single kidney transplanted patients for past history, graft surgery, basal donor and recipient demographic characteristics and immunosuppression treatment. We also decided to restrict the inclusion to first and second transplantations and to transplantations from a deceased donor, because surveillance biopsies are not routinely performed in recipients with three or more transplantations and with transplantation from a living donor, the risk of adverse events related to surveillance biopsy being ethically questionable. BK virus replication was measured from patient blood samples in each participating center within the first year post-transplantation at 3, 6 and 12 months. Patients with active replication were not included in the study as a kidney biopsy is usually performed to eliminate a diagnosis of BK virus nephropathy and minimization of the immunosuppression is indicated.

For training and internal validation, data were extracted from the French DIVAT cohort of kidney transplant recipients (www.divat.fr, final agreement from the French Commission of the CNIL, decision DR-2025-087, February 15, 2015). Only the University Hospital of Lyon, Paris Necker and Nantes were included as the four other centers belonging to the DIVAT network did not perform 1-year surveillance biopsies. We limited the training cohort to the 466 recipients transplanted between January 2006 and June 2012 since the 1-year surveillance biopsies were concomitantly used between the three centers within this period. Among the training cohort, 453 patients (97.2%) received CNI as maintenance therapy with 71.9% of patients under tacrolimus. Almost all patients (97%) were treated with mycophenolic acid derivatives (82.8% under mycophenolate mofetil [MMF] and 14.2% under MPA) and 92.3% of patients received a corticosteroid regimen. For external validation, we merged together two other cohorts: (i) a subgroup of recipients of the DIVAT cohort enrolled in a biomarker research study entitled VALBIO12 (ethical comity #PROG/11/48, July 9, 2013) constituted with 174 patients not included in the training set, and (ii) 545 patients from a Belgian cohort (Leuven University Hospital, #S53364 Biobank Renal Transplantation). As for the validation cohort, the same inclusion criteria were used.

Transplant International 2018; 31: 947–955

Available data

The following data relative to the donor were extracted: age, gender, last serum creatinine and cause of death. Recipient features were: age, gender, Body Mass Index (BMI), history of hypertension, cardiovascular disease, diabetes, dyslipidemia and/or obesity, cause of initial renal disease (recurrent nephropathy versus other, detailed in Table S1), transplantation rank. Pretransplantation immunization against Anti Human Leukocyte Antigen (HLA) was defined as positive if at least one DSA was identified by Luminex® Single Antigen Beads Technology (One Lambda, Los Angeles, CA, USA for Nantes and Necker and GenProb® USA for Lyon) or if Luminex® screening or another technology (ELISA or CDC) was positive if Luminex® Single Antigen Beads was not performed within the 6 months before the transplantation. Transplantation parameters were cold ischemia time, number of HLA-A-B-DR incompatibilities, Delayed Graft Function (DGF, defined as the need for dialysis after transplantation), and serum creatinine at 3, 6, and 12 month post-transplantation. Biopsy proven acute rejection, cytomegalovirus (CMV) disease, and graft acute pyelonephritis were collected within the first year after the transplantation prior the 1-year biopsy.

Definition of endpoint

The objective of our study was to propose a diagnostic test for abnormal histology (AH) on the 1-year surveillance biopsy. We therefore proposed defining two groups of patients based on histological diagnosis (normal or subnormal histology [NSH] versus abnormal histology). We assumed that isolated and noninflammatory mild IFTA grade 1 are nonsevere lesions. We considered these lesions as subnormal histology and pooled these with normal histology. Biopsies presenting an isolated "i" score at 1 were considered as normal histology. Patients with normal histology or mild IFTA grade 1 or isolated allograft glomerulopathy "cg" with no C4d, no microvascular inflammation, and no DSA were pooled in a single group of NSH. The patients with abnormal histology were those who displayed severe IFTA (grades 2 and 3) with or without inflammation, allograft rejection (acute or chronic T-cell mediated rejection and acute or chronic antibody-mediated rejection, including borderline changes) and recurrent or *de novo* glomerulonephritis other than allograft glomerulopathy "cg." The individual Banff scores on the 1-year surveillance biopsies were prospectively performed by each transplantation center in addition to anti-HLA DSA identification at the time of the 1-year surveillance biopsy. Our local pathologist (K.R.) retrospectively centralized and re-classified each patient into one of the two groups (NSH or AH) according to the last Banff 2013 criteria [16]. These elementary lesions given the histological diagnostics (Normal, IFTA 1 and isolated cg for the NSH group and rejection, IFTA 2 and 3 and glomerulonephritis for the AH group) are presented on Fig. S1.

Statistical analysis

Comparisons of characteristics regarding both AH and NSH groups were performed using Student tests or chi-square test (eventually Fisher exact test when appropriate), respectively for quantitative or categorical variables. The 1-year Renal Biopsy Index (1-RBI) was the linear predictor of a logistic regression. Quantitative variables were possibly categorized according to clinically relevant thresholds if log-linearity was not graphically verified. Relevant clinical interactions between explanatory variables were also tested, such as the interactions between serum creatinine at 3, 6, or 12 months and recipient or donor gender. The selection of explanatory variables in the 1-RBI was performed using a Lasso penalty, which is a convenient method to select a sparse model faced with numerous explanatory variables [17], the corresponding tuning parameter was estimated by 5-fold cross-validation. The 0.632+ bootstrap estimator of the ROC curve was used for internal validation of discriminative capacities [18], while external validation was performed by estimating usual ROC curve, the 95% Confidence Interval (95% CI) being nonparametrically computed by bootstrap resampling. The calibration, that is the concordance between the observed AH probabilities and the expected ones, was performed from 12 intervals and the Hosmer–Lemeshow statistic.

All analyses were performed using the 3.2.0. version of the R software [19]. The ROC632 package (version 0.6) was used for implementing the logistic regression with a Lasso penalization and the 0.632+ algorithm.

Results

Characteristics of training cohort

Among the 466 patients of the training cohort, 229 recipients (49.1%) were diagnosed with abnormal histology: 132 patients with rejection, 59 patients with

IFTA grade 2, 26 with IFTA grade 3 and 12 with glomerulonephritis. The normal or sub-normal histology (NSH) group was constituted by 237 patients (50.9%) including 88 patients with normal histology (including seven biopsies with isolated "i" score at 1), 141 with IFTA grade 1 without inflammation in the nonscarred "i" and total area "ti" and 8 with isolated "cg." The characteristics of the cohort are described in Table 1. The mean recipient age was 48.8 years (±12.5), 61.8% were men, and 81.6% were recipients of a first kidney transplant, while 28.8% were recipients with a potential recurrent initial disease. The mean donor age was 49.6 years (±16.2) and 61.6% were male, including half who died of vascular brain damage.

Description of 1-year Renal Biopsy Index

Five variables were retained in the score (Table 2). Female recipients (OR = 3.3001, 95% CI from 2.0247 to 5.4526), patients with pretransplantation anticlass II immunization (OR = 1.7748, 95% CI from 1.1402 to 2.7888) and increased serum creatinine levels at 3 months (OR = 1.0028, 95% CI from 0.9955 to 1.0103; let us recall that the 95% CI may include value 1 as variables were not retained on their significant association but on their predictive abilities), 6 months (OR = 1.0083, 95% CI from 0.9990 to 1.0178) and 12 months (OR = 1.0082, 95% CI from 0.9994 to 1.0173) were factors retained for their contribution to the prediction of the AH probability. The 1-RBI can be calculated by summing up the OR logarithm multiplied by the values of the explanatory variables:

$$\begin{aligned}\text{1-RBI} = {} & \log(1.7748) \times [1 \text{ if (Positive Anticlass II immunization) and 0 otherwise}] \\ & + \log(0.3030) \times [1 \text{ if (Male Recipient gender) and 0 otherwise}] \\ & + \log(1.0028) \times [\text{Recipient serum creatinine at 3 months in } \mu\text{mol/l}] \\ & + \log(1.0083) \times [\text{Recipient serum creatinine at 6 months in } \mu\text{mol/l}] \\ & + \log(1.0082) \times [\text{Recipient serum creatinine at 12 months in } \mu\text{mol/l}]\end{aligned}$$

Faced with the collinearity of serum creatinine measures at 3, 6, and 12 months, we also intended to replace it by the difference of the serum of creatinine between two times. However, we did not achieve better predictive performance. Finally, the discriminative capacities of the 1-RBI corresponded to an area under ROC curve (AUC) at 0.71 (internal validation, 95% CI from 0.53 to 0.83), meaning that we have a 71% chance of observing a score higher in a recipient with AH compared to another with NSH. We hypothesized that (i) centers where no surveillance biopsy is performed will accept this invasive examination for patients with at least a 80% risk of AH (positive predictive value, PPV), and (ii) centers where surveillance biopsies are performed will accept the absence of examination if a patient had at least 80% risk of NSH (negative predictive value, NPV). The PPV at 80% corresponds to define a positive test when the 1-RBI value is higher than 2.81 (Fig. 1). The corresponding NPV was 58%. In our cohort of 427 patients without missing data on the 1-RBI, 63 patients (15%) had 1-RBI higher than 2.81. Among these 63 patients, 53 presented AH (27 rejections, 12 IFTA 2, 11 IFTA 3 and three recurrent glomerulonephritis), while 10 patients were misclassified as presenting normal or IFTA grade 1. The NPV at 80% corresponds to define a negative test when 1-RBI value is lower than 0.43. But among the 427 patients, only three patients (1%) had 1-RBI lower than 0.43, demonstrating the incapacity of the 1-RBI to discriminate patients with such a high chance of NSH.

External validation

Among the 647 independent French and Belgian recipients without missing data on variables constituting the 1-RBI, 326 (50.4%) were diagnosed with AH, a prevalence comparable to the one observed in the training cohort. As presented in Table S2, we note that some characteristics seemed to be different between the training and validation cohorts. For instance, the validation cohort included less secondary grafts (9.9% vs. 18.5% in the training cohort) or had less frequent DGF (19.8% vs. 32% in the training cohort). The discriminative capacities of the 1-RBI were associated with an AUC at 0.62 (95% CI from 0.58 to 0.66). The illustrated calibration plot (Fig. 2) predicting probabilities of AH were significantly under-estimated using the 1-RBI ($P < 0.0001$, Hosmer–Lemeshow statistic). As a consequence, the PPV for a threshold at 2.81 was 70%. The corresponding NPV was 54%. Around 17% of recipients displayed a 1-RBI value higher than 2.81, which was comparable to the percentage estimated in the training cohort. In a population presenting a similar prevalence of abnormal histology, we may reasonably conclude that we validated the proposed decision rule.

Transplant International 2018; 31: 947–955

Table 1. Description of recipient, donor, and renal transplant characteristics at time of transplantation and within the first year after the transplantation on the training cohort and according to lesion groups (abnormal histology versus normal or subnormal histology).

	Missing data	Global *N* = 466	AH *N* = 229 (49.1%)	NSH *N* = 237 (50.9%)	*P*-value
Quantitative characteristics: mean ± SD					
Recipient age (years)	0	48.83 ± 12.47	48.96 ± 12.87	48.70 ± 12.08	0.8200
HLA incompatibilities ABDR	0	3.06 ± 1.33	3.14 ± 1.33	2.99 ± 1.33	0.2157
Recipient BMI (kg/m^2)	0	23.67 ± 4.13	23.88 ± 4.36	23.47 ± 3.90	0.2834
Recipient SCr at 3 months (μmol/l)	2	134.92 ± 53.84	144.86 ± 63.20	125.39 ± 40.92	0.0001
Recipient SCr at 6 months (μmol/l)	7	132.95 ± 46.96	143.42 ± 53.80	122.79 ± 36.54	<0.0001
Recipient SCr at 12 months (μmol/l)	3	132.70 ± 47.12	143.72 ± 53.04	122.19 ± 37.91	<0.0001
Donor age (years)	0	49.57 ± 16.18	51.03 ± 15.82	48.16 ± 16.43	0.0557
Donor SCr (μmol/l)	0	97.80 ± 73.70	94.70 ± 58.96	100.79 ± 85.58	0.3708
Cold ischemia time (h)	0	20.29 ± 7.10	20.32 ± 7.18	20.26 ± 7.04	0.9209
Categorical characteristics: *N* (%)					
Recipient men	0	288 (61.80)	126 (55.02)	162 (68.35)	0.0042
Glomerulonephritis	0	134 (28.76)	65 (28.38)	69 (29.11)	0.9429
Anticlass I immunization >0	23	147 (31.55)	83 (36.24)	64 (27.00)	0.0243
Anticlass II immunization >0	28	143 (30.69)	84 (36.68)	59 (24.89)	0.0036
History of hypertension	0	390 (83.69)	189 (82.53)	201 (84.81)	0.5893
History of cardiovascular disease	0	90 (19.31)	40 (17.47)	50 (21.10)	0.3816
Recipient History of type I diabetes	0	25 (5.36)	18 (7.86)	7 (2.95)	0.0320
Recipient History of type II diabetes	0	30 (6.44)	14 (6.11)	16 (6.75)	0.9271
History of dyslipidemia	0	121 (25.97)	60 (26.20)	61 (25.74)	0.9935
History of obesity	0	55 (11.80)	25 (10.92)	30 (12.66)	0.6608
Cerebrovascular donor death	1	252 (54.08)	127 (55.46)	125 (52.74)	0.6554
Donor men	0	287 (61.59)	147 (64.19)	140 (59.07)	0.2979
Second transplantation	0	86 (18.45)	50 (21.83)	36 (15.19)	0.0838
Depleting induction	0	184 (39.48)	109 (47.6)	156 (34.65)	0.0006
DGF	0	149 (31.97)	72 (31.44)	77 (32.49)	0.8861
CMV infection prior 1-year biopsy	0	43 (9.23)	28 (12.23)	15 (6.33)	0.0414
Acute graft pyelonephritis prior 1-year biopsy	0	53 (11.37)	37 (16.16)	16 (6.75)	0.0023
Cystitis prior 1-year biopsy	0	154 (32.05)	79 (34.50)	75 (31.65)	0.5783
Acute rejection prior 1-year biopsy					
Cellular	0	52 (11.16)	35 (15.28)	17 (7.17)	0.0046
Humoral		8 (1.7)	6 (2.62)	2 (0.84)	

Table 2. Results of the lasso penalized logistic regression (n = 427, 39 observations removed due to missing data).

	Log(OR)	OR	95% CI
Recipient Scr at 3 months (µmol/l)	0.0028	1.0028	0.9955–1.0103
Recipient SCr at 6 months (µmol/l)	0.0083	1.0083	0.9990–1.0178
Recipient SCr at 12 months (µmol/l)	0.0082	1.0082	0.9994–1.0173
Recipient men (men versus women)	−1.1940	0.3030	0.1834–0.4939
Anticlass II immunization (positive versus negative)	0.5737	1.7748	1.1402–2.7888

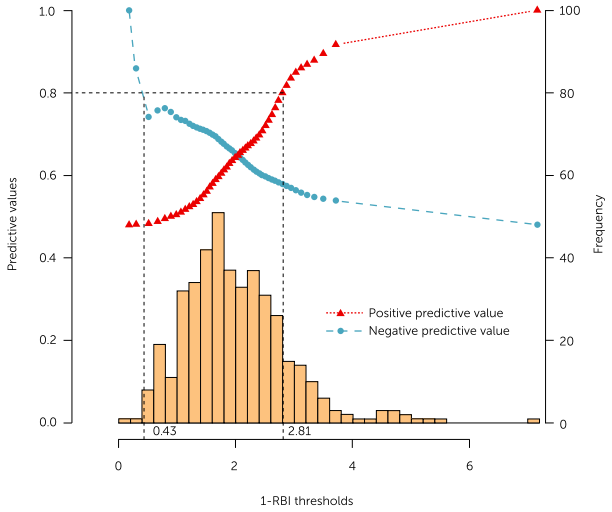

Figure 1 Positive and negative predictive values according to the possible 1-year Renal Biopsy Index thresholds (n = 427, 39 observations removed due to missing data).

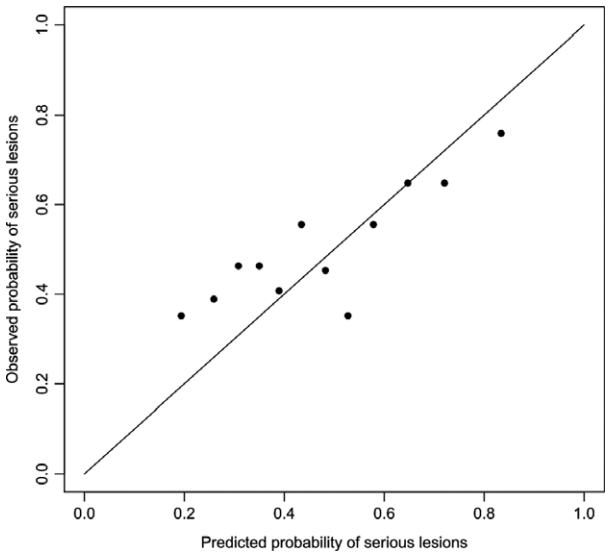

Figure 2 Evaluation of the calibration of the 1-year Renal Biopsy Index from the external validation sample (n = 647, 72 observations removed due to missing data). The predicted and observed probabilities of abnormal histology were calculated for 12 intervals ($P < 0.0001$, Hosmer–Lemeshow statistic).

About two examples to illustrate the 1-RBI usefulness in clinical practice

Let consider a first example of a non-anti-HLA immunized 65-year-old woman receiving a first kidney from a 70-years-old cerebrovascular dead donor with a creatinemia 66 µmol/l, with 11 h of cold ischemia time, presenting a serum creatinine at 151 µmol/l at 3 months, 147 µmol/l at 6 months and 145 µmol/l at 12 months, a 1 year proteinuria at 0.35 g/day, no rejection observed during the first year. The corresponding 1-RBI calculation was 2.84, meaning that she could present an 80% risk of displaying abnormal histology on her 1-year biopsy. For this patient, we actually observed an infraclinic rejection from the 1-year surveillance biopsy.

As a second example, we considered a non-anti-HLA immunized 37-year-old woman receiving a first kidney from a 50-years-old noncerebrovascular dead donor with a creatinemia 163 µmol/l, with 20 h of cold ischemia time, presenting a serum creatinine at 132 µmol/l at 3 months, 156 µmol/l at 6 months and 150 µmol/l at 12 months, without proteinuria, no rejection observed during the first year. The corresponding 1-RBI calculation was 2.90, meaning that she could have an 80% risk of displaying abnormal histology on her 1-year biopsy. Finally, the result of biopsy also showed an infraclinic rejection.

Discussion

The choice of performing surveillance biopsy at 1-year post-transplantation is debatable. Physicians have heterogeneous policies regarding the absence of well-established guidelines. We therefore developed a clinical-based diagnostic tool of abnormal histological lesions on the 1-year surveillance biopsy to help physicians in their decision to perform this invasive examination among patients alive with a functioning graft at 1-year post-transplantation.

This scoring system, named 1-RBI for 1-year Renal Biopsy Index, is based on 5 variables available in the routine: the recipient gender, pretransplantation anti-class II immunization, and serum creatinine levels at 3, 6, and 12 months. We demonstrated acceptable discriminative capacities from internal validation (AUC = 0.71, 95% CI from 0.53 to 0.83), which were slightly deteriorated on the external validation (AUC = 0.62, 95% CI from 0.58 to 0.66). Beyond the intrinsic discriminative capacities of the 1-RBI, its clinical relevance relies on a required low rate of error when recommending a 1-year biopsy due to a high 1-RBI, that is PPV at 80%. This medical decision tool could specifically help physicians who do not routinely practice 1-year surveillance biopsies and allow them to not miss potential actionable lesions despite there being no alarming clinical or biological parameters at 1-year of follow-up. Our study included patients of transplantation centers having a 1-year surveillance biopsy policy on which we described 15% of patients presenting a 1-RBI higher than 2.81 for whom a 1-year biopsy could be recommended. Besides, among 760 patients from DIVAT transplantation centers without a 1-year surveillance biopsy policy and following the same inclusion criteria, we could estimate that 13% of patients would present a 1-RBI higher than 2.81 for whom a 1-year biopsy could be recommended while they did not present suspicious clinical signs. In contrast, we could not propose a decision rule to convince transplantation centers usually performing a 1-year surveillance biopsy to avoid biopsies for patients on the basis of 1-RBI calculation as only 1% of patients presented a high chance of normal or mild IFTA for a NPV of 80%, and a corresponding threshold of 1-RBI of 0.43. The 1-RBI seems to not be a surrogate for avoiding a surveillance biopsy, but instead a simple clinical tool to help indication for a 1-year biopsy when there is a good likelihood of observing potential actionable histological lesions despite bleeding risk after a biopsy and despite being aware that there is no clear therapeutic consensus according to the identification of most histological lesions [20,21]. Note that we did not observe abnormal bleeding rate compared to the literature with 1% of hematoma and serious bleeding and 0.05% isolated and rapidly recovering hematuria since 2006 in our whole cohort [22,23]. Many transplantation centers that do not routinely practice a 1-year surveillance biopsy would perform a biopsy for high serum creatinine levels or a significant increase in serum creatinine within the first year after transplantation. We think that our proposed scoring system could benefit in case of intermediate clinical situations: where it is difficult to appreciate the risk of abnormal histology while there is no obvious clinical or biological signs, thus encouraging physicians to predict a risk of abnormal histology without delay for potential actionable lesions to treat. From a patient-centered point of view, this scoring system could also help patients and physicians for shared medical decision about a 1-year biopsy indication despite stable graft function.

Various limitations of our study have to be underlined. Firstly, the 1-RBI may appear not convincing enough for the physicians who currently performed 1-year systematic biopsy to abandon this examination. We believe that the discriminative capacities of the 1-RBI can be improved by adding, for instance, the 3 and 6 month daily proteinuria that were not included into the 1-RB1 due to numerous missing data in our cohort. Nevertheless, one can argue that if patients present a significant and confirmed proteinuria, physician will indicated a causal biopsy. Besides, Rabant *et al.* [24] showed that the urinary CXCL10:Cr ratio at 3 months post-transplantation could predict immunological quiescence on a triple-drug CNI-based immunosuppressive regimen in clinically and histologically stable patients at 1-year post-transplantation. Thus, the inclusion of such a marker in the 1-RBI could lead to achieving a diagnostic tool that is more accurate in terms of NPV. There is also an increasing interest in other markers that could predict histological lesions, such as acute rejection [20,21,24]. Secondly, the definition of both AH and NSH groups can be discussed. This NSH group is composed by patients who presented either normal histology or IFTA grade 1. We made this choice because there is currently no evidence of benefits from therapeutic intervention for these histological features. Also, the AH group gathered either allo-immune (including borderline lesions) or severe lesions of IFTA grade 2 or 3 with or without inflammation in the scarred or non-scarred area or glomerulonephritis. Our choice was made in accordance with the one made by Cosio *et al.* [13] who divided patients into two groups: one gathered normal histology, IFTA grade 1 or interstitial inflammation without IFTA compared and one pooled all the other lesions including glomerulonephritis other than transplant glomerulopathy. We were not able for logistic and financial reasons to afford a centralized re-reading of histology slides by an independent pathologist despite it would have reduced heterogeneity of Banff scoring due to disagreement between pathologists of each center. However to limit this bias, we made the choice to retrospectively re-classify all histological

diagnosis from the elementary lesions description according to the Banff criteria thanks to our local pathologist (K.R.) [16]. In our training cohort, one of five patients presented normal histology and 1/3 an IFTA grade 1 that represented more than half of the histological features. Pooling humoral rejection with cellular rejection and severe IFTA could be arguable. However, our choice was to merge these lesions together as we thought that there were more therapeutic options for these patients. Nevertheless, we were aware that there is no consensus to modify therapy according to biopsy findings except for subclinical rejection [11,13]. In addition, Cosio *et al.* [13] showed that patients with normal histology have a high probability to maintained benign lesions after 1-year is consistent with the classification of these patients in the NSH group.

Finally, as already mentioned, the 1-RBI presented a reasonable discriminative power (AUC = 0.71, 95% CI from 0.53 to 0.83), probably good enough to provide a medical decision making tool especially for physicians who do not routinely practice a 1-year surveillance biopsy. Indeed, for a PPV defined at 80%, it was possible to define a 1-RBI threshold to screen patients at high risk of presenting abnormal histology. We confirmed the discriminative power of the 1-RBI in an independent external validation cohort, but nevertheless European, showing similar demographic parameters. As predictive values depend on the prevalence of abnormal histology, it is possible that the proposed stratified decision rule to indicate a 1-year biopsy could not apply for US patients, such as those belonging to the Cosio's cohort recently published [13], due to the different prevalence of "unfavorable lesions" notably induced by a different rate of acute rejection. Cosio's definition of "unfavorable lesions" differed slightly from our definition of "abnormal histology." Indeed, we considered as abnormal histology all types of rejection (Acute or chronic TCMR and Acute or chronic ABMR) as well as borderline changes, Cosio considered only cAMR while 4% of TCMR was split between the "favorable group" (acute inflammation without fibrosis) and the "unfavorable group" (IFTA + i). Another explanation could be an indication bias in patients with renal instability. Additionally, differences in demographic features between both cohorts may have accentuated the difficultly in generalizing 1-RBI use. For instance, 78% of patients in Cosio's cohort received living donor kidneys while we excluded them from our study. Mean donor age was 45.2 years in Cosio's cohort with mainly female donors (54.3%) compared to 49.6 years old and 38.4% female donors in our cohort. As mentioned in the methodological and epidemiological literature, in such a situation where predictive values would not be robust, it may be necessary to recalibrate the model and eventually to estimate a new prediction model that could include specific risk factors [25–27]. The demographic characteristics of our cohort are close to those of a French multicenter retrospective study [9] in which the rate of "abnormal histology" pooling IFTA II/III, rejection and glomerulonephritis was 37.3%, in between Cosio's cohort and our recent study.

In conclusion, we proposed the 1-RBI, a clinical score that may be useful for a more standardized proposal of 1-year biopsy. An online calculator is available at www.divat.fr. Beyond the complete description of its discriminative capacities between recipients with and without abnormal lesions, we proposed a test that could help physicians, especially those who do not currently perform surveillance biopsy. More precisely, this score may for instance help to indicate a 1-year biopsy in patients without suspicious clinical or biological signs but presenting a 1-RBI level higher than 2.81. But the discriminative capacities of the 1-RBI have to be improved, for instance by including biomarkers, especially to achieve a better NPV for further recommendations in low-risk patients. Other studies are also needed to propose efficient therapeutics according to biopsy results.

Funding

This work was supported by the French National Research Agency [ANR-11-JSV1-0008-01, 2011] and the French Biomedicine Agency. We thank the Roche Pharma, Novartis, and Sanofi laboratories for supporting the DIVAT Cohort as the CENTAURE Foundation (http://www.fondation-centaure.org).

Conflict of interest

The authors declare that there is no conflict of interest with Roche Pharma, Novartis, and Sanofi, which partially supported DIVAT cohort funding.

SUPPORTING INFORMATION

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

Figure S1. Description of each elementary lesion given the histological diagnostics (Normal, IFTA 1 and

isolated cg for the NSH group and rejection, IFTA 2 and 3 and glomerulonephritis for the AH group).

Table S1. Detailed description of the cause of initial renal disease on the training cohort and according to lesion groups (AH versus NSH).

Table S2. Description of recipient, donor and renal transplant characteristics at time of transplantation and within the first year after the transplantation on the validation cohort and according to lesion groups (AH versus NSH).

REFERENCES

1. Yilmaz S, Tomlanovich S, Mathew T, *et al.* Protocol core needle biopsy and histologic Chronic Allograft Damage Index (CADI) as surrogate end point for long-term graft survival in multicenter studies. *J Am Soc Nephrol* 2003; **14**: 773.
2. Nankivell BJ, Borrows RJ, Fung CL-S, O'Connell PJ, Allen RDM, Chapman JR. The natural history of chronic allograft nephropathy. *N Engl J Med* 2003; **349**: 2326.
3. Hertig A, Anglicheau D, Verine J, *et al.* Early epithelial phenotypic changes predict graft fibrosis. *J Am Soc Nephrol* 2008; **19**: 1584.
4. Cosio FG, Grande JP, Wadei H, Larson TS, Griffin MD, Stegall MD. Predicting subsequent decline in kidney allograft function from early surveillance biopsies. *Am J Transplant* 2005; **5**: 2464.
5. Lefaucheur C, Nochy D, Bariety J. Renal biopsy: procedures, contraindications, complications. *Néphrol Thér* 2009; **5**: 331.
6. Rush D, Arlen D, Boucher A, *et al.* Lack of benefit of early protocol biopsies in renal transplant patients receiving TAC and MMF: a randomized study. *Am J Transplant* 2007; **7**: 2538.
7. Kurtkoti J, Sakhuja V, Sud K, *et al.* The utility of 1- and 3-month protocol biopsies on renal allograft function: a randomized controlled study. *Am J Transplant* 2008; **8**: 317.
8. Kee TY-S, Chapman JR, O'Connell PJ, *et al.* Treatment of subclinical rejection diagnosed by protocol biopsy of kidney transplants. *Transplantation* 2006; **82**: 36.
9. Moulin B, Merville P, Renaudin K, *et al.* Evaluation of protocol biopsy utility 12 months after renal transplantation: a multicenter observational analysis. *J Transplant* 2012; **2012**: 781263.
10. Sis B, Jhangri GS, Riopel J, *et al.* A new diagnostic algorithm for antibody-mediated microcirculation inflammation in kidney transplants. *Am J Transplant* 2012; **12**: 1168.
11. Loupy A, Vernerey D, Tinel C, *et al.* Subclinical rejection phenotypes at 1 year post-transplant and outcome of kidney allografts. *J Am Soc Nephrol* 2015; **26**: 1721.
12. Stegall MD, Park WD, Larson TS, *et al.* The histology of solitary renal allografts at 1 and 5 years after transplantation. *Am J Transplant* 2011; **11**: 698.
13. Cosio FG, El Ters M, Cornell LD, Schinstock CA, Stegall MD. Changing kidney allograft histology early posttransplant: prognostic implications of 1-year protocol biopsies. *Am J Transplant* 2015; **16**: 194.
14. de Sandes-Freitas TV, Felipe CR, Campos ÉF, *et al.* Subclinical lesions and donor-specific antibodies in kidney transplant recipients receiving tacrolimus-based immunosuppressive regimen followed by early conversion to sirolimus. *Transplantation* 2015; **99**: 2372.
15. Servais A, Meas-Yedid V, Noël LH, *et al.* Interstitial fibrosis evolution on early sequential screening renal allograft biopsies using quantitative image analysis. *Am J Transplant* 2011; **11**: 1456.
16. Solez K, Racusen LC. The Banff classification revisited. *Kidney Int* 2013; **83**: 201.
17. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. *J R Stat Soc Ser B Stat Methodol* 2011; **73**: 273.
18. Foucher Y, Danger R. Time dependent ROC Curves for the estimation of true prognostic capacity of microarray data. *Statistical Applications in Genetics and Molecular Biology*. 2012; **11**: Article 1. DOI:10.1515/1544-6115.1815.
19. R Development Core Team. *R: A Language and Environment for Statistical Computing* [Internet]. Computing RF for S, editor. Vienna, Austria: R Development Core Team, 2010. Available from: http://www.R-project.org/.
20. Simmonds MJ. Using genetic variation to predict and extend long-term kidney transplant function. *Transplantation* 2015; **99**: 2038.
21. Ho J, Rush DN, Karpinski M, *et al.* Validation of urinary CXCL10 as a marker of borderline, subclinical, and clinical tubulitis. *Transplantation* 2011; **92**: 878.
22. Hergesell O, Felten H, Andrassy K, Kühn K, Ritz E. Safety of ultrasound-guided percutaneous renal biopsy-retrospective analysis of 1090 consecutive cases. *Nephrol Dial Transplant* 1998; **13**: 975.
23. Furness PN, Philpott CM, Chorbadjian MT, *et al.* Protocol biopsy of the stable renal transplant: a multicenter study of methods and complication rates. *Transplantation* 2003; **76**: 969.
24. Rabant M, Amrouche L, Lebreton X, *et al.* Urinary C-X-C motif chemokine 10 independently improves the noninvasive diagnosis of antibody-mediated kidney allograft rejection. *J Am Soc Nephrol* 2015; **26**: 2840.
25. Steyerberg EW, Moons KGM, van der Windt DA, *et al.* Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 2013; **10**: e1001381.
26. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* 2009; **338**: b606.
27. Moons KGM, Altman DG, Reitsma JB, *et al.* Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015; **162**: W1.