UCLA
UCLA Previously Published Works

Title

Validation of a prediction system for risk of kidney allograft failure in pediatric kidney transplant recipients: An international observational study.

Permalink

https://escholarship.org/uc/item/3f50s3t2

Journal

American Journal of Transplantation, 23(10)

Authors

Hogan, Julien
Divard, Gillian
Aubert, Olivier
et al.

Publication Date

2023-10-01

DOI

10.1016/j.ajt.2023.07.004

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Am J Transplant. Author manuscript; available in PMC 2024 October 01.

Published in final edited form as:

Am J Transplant. 2023 October ; 23(10): 1561–1569. doi:10.1016/j.ajt.2023.07.004.

Validation of a prediction system for risk of kidney allograft failure in pediatric kidney transplant recipients: An international observational study

Julien Hogan[1,2,3,†], **Gillian Divard**[1,†], **Olivier Aubert**[1], **Rouba Garro**[4], **Olivia Boyer**[5], **Lee Alex Donald Cooper**[6], **Alton Brad Farris**[7], **Marc Fila**[8], **Michael Seifert**[9], **Anne-Laure Sellier-Leclerc**[10], **Jody Smith**[11], **Alexander Fichtner**[12], **Burkhard Tönshoff**[12], **Katherine Twombley**[13], **Bradley Warady**[14], **Meghan Pearl**[15], **Rima S. Zahr**[16], **Carmen Lefaucheur**[1], **Rachel Patzer**[3], **Alexandre Loupy**[1,*]

[1]Université Paris Cité, INSERM, UMR-S970, PARCC, Paris Translational Research Center for Organ Transplantation, Paris, France

[2]Pediatric nephrology department, Robert Debreé Hospital, APHP, Paris, France

[3]Emory Transplant Center, Department of Surgery, Emory University, Atlanta, Georgia, USA

[4]Pediatric Nephrology Department, Children Healthcare of Atlanta, Emory University, Atlanta, Georgia, USA

[5]Pediatric Nephrology, MARHEA Reference Center, INSERM U1163, Imagine Institute, Paris Cité University, Necker-Enfants Malades Hospital, APHP. Centre, Paris, France

[6]Department of Pathology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

[7]Department of Pathology and Laboratory Medicine, Emory University School of Medicine, Atlanta, Georgia, USA

[8]Pediatric Nephrology Department, Montpellier University Hospital, Montpellier, France

[9]Pediatric Nephrology Department, University of Alabama, Birmingham, Alabama, USA

[10]Pediatric Nephrology Department, Mother and Child University Hospital, HCL, Lyon, France

[11]Pediatric Nephrology Department, Seattle Children, Seattle, New York, USA

[12]Department of Pediatrics I, University Childrens Hospital Heidelberg, Heidelberg, Germany

[13]Pediatric Nephrology Department, Medical University of South Carolina, Charleston, South Carolina, USA

[*]Corresponding author. Paris Translational Research Center for Organ Transplantation, 56 Rue Leblanc, Paris 75015, France., alexandre.loupy@inserm.fr (A. Loupy).

[†]These authors contributed equally: Julien Hogan and Gillian Divard.

Disclosure

The authors of this manuscript have no conflicts of interest to disclose as described by *the American Journal of Transplantation*.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajt.2023.07.004.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[14]Pediatric Nephrology Department, Children's Mercy, Kansas City, Michigan, USA

[15]Pediatric Nephrology Department, David Geffen School of Medicine at UCLA, Los Angeles, California, USA

[16]UTHSC Department of Pediatric Nephrology and Hypertension, Le Bonheur Children's Hospital, Memphis, Tennessee, USA

Abstract

Predicting long-term kidney allograft failure is an unmet need for clinical care and clinical trial optimization in children. We aimed to validate a kidney allograft failure risk prediction system in a large international cohort of pediatric kidney transplant recipients. Patients from 20 centers in Europe and the United States, transplanted between 2004 and 2017, were included. Allograft assessment included estimated glomerular filtration rate, urine protein-to-creatinine ratio, circulating antihuman leukocyte antigen donor-specific antibody, and kidney allograft histology. Individual predictions of allograft failure were calculated using the integrative box (iBox) system. Prediction performances were assessed using discrimination and calibration. The allograft evaluations were performed in 706 kidney transplant recipients at a median time of 9.1 (interquartile range, 3.3–19.2) months posttransplant; mean estimated glomerular filtration rate was 68.7 ± 28.1 mL/min/1.73 m^2, and median urine protein-to-creatinine ratio was 0.1 (0.0–0.4) g/g, and 134 (19.0%) patients had antihuman leukocyte antigen donor-specific antibodies. The iBox exhibited accurate calibration and discrimination for predicting the outcomes up to 10 years after evaluation, with a C-index of 0.81 (95% confidence interval, 0.75–0.87). This study confirms the generalizability of the iBox to predict long-term kidney allograft failure in children, with performances similar to those reported in adults. These results support the use of the iBox to improve patient monitoring and facilitate clinical trials in children.

Keywords

kidney transplantation; children; predictive model; allograft failure; Banff classification

1. Introduction

Despite major improvement in kidney allograft survival in children over the last decades, most pediatric kidney transplant (kTx) recipients will experience graft failure and require additional kidney transplantation with or without periods of dialysis. Although kidney allograft survival is high in this population, with 10 years' survival of 73.2% in Europe[1] and between 65% and 75% in the United States depending on the type of donor,[2] graft survival varies greatly between patients, and its accurate prediction is essential to inform patient care (prepare retransplantation or anticipate dialysis access creation), improve risk stratification, and design more pragmatic clinical trials. There is currently no accurate predictive model of long-term kidney allograft survival validated among pediatric kTx recipients. Some previous prediction models included adolescents in the development of adult models.[3] Krikov et al[4] included data from both pediatric and adult patients available in the United States Data Registry (USRDS), but the model was not specifically developed or validated in the pediatric population and was lacking important and independent predictors of allograft survival such

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

as the presence and intensity of donor-specific antibodies (DSA) or histological findings on allograft biopsies. Finally, 2 more recent studies from the ESPN/ERA-EDTA registry and from the French REIN registry developed prediction models specifically in the pediatric population.[5,6] Both studies reported good discrimination, but none of these studies were externally validated, leaving their generalizability in doubt and preventing their use as clinical decision tools or valid candidates for predicting hard endpoints in the setting of clinical trials. Recently, the integrative box risk prediction (iBox) prediction system[7] was developed by a consortium of 15 clinical transplant centers in Europe and in the US to meet the need for accurate prediction of long-term renal allograft failure. The scoring system was derived from clinical, biological, immunological, and histological parameters prospectively measured as part of standard of care during kTx patient follow-up and further validated in 6 randomized controlled trials (RCTs), suggesting that this score could be used as a surrogate endpoint in clinical trials.[8,9] This score was formally accepted into the Biomarker Qualification Program of the Food and Drug Administration,[10] the first step toward formal regulatory endorsement, and was recently qualified as an efficacy endpoint in clinical trials investigating novel immunosuppressive medicines in kTx patients by the European Medicines Agency.[11] We hypothesized that the multimodality approach of the iBox scoring system could allow its generalizability and transferability to pediatric transplant recipients. The aim of this study was to assess the accuracy of the iBox score in a large, international, and well-phenotyped cohort of pediatric kTx recipients.

2. Patients and methods

2.1. Patients

The pediatric validation cohort included recipients of kTx from a living or a deceased donor who were <21 years of age at the time of transplantation and were transplanted between January 1, 2004, and December 31, 2017. All 1359 consecutive children who received a kTx over the study period in one of the participating centers were screened for inclusion. We excluded patients with primary nonfunctioning transplants and/or who did not have a kidney allograft biopsy performed during follow-up. A flowchart of the study population and a comparison of the patients included with those not included are provided in Supplementary Table and Supplementary Figure S1. Participating centers included 20 centers from Europe and the United States. The European centers were Hôpital Robert Debré, Paris, France; Hôpital Necker Enfants Malades, Paris, France; Hôpital Mére-Enfant, Lyon, France; Hôpital Arnaud de Villeneuve, Montpellier, France; and 10 other European centers providing data to the Cooperative European Paediatric Renal Transplant Initiative (CERTAIN) registry[12] (Cologne, Essen, Hamburg, Heidelberg, Münster, Tübingen [Germany]; Dublin [Ireland]; Padua, Rome [Italy]; Zürich [Switzerland]). The US centers were Children's Healthcare of Atlanta, Atlanta, Georgia; University of Alabama, Birmingham, Albama; Seattle Children, Seattle, New York; Medical University of South Carolina, Charleston, South Carolina; Children's Mercy, Kansas City, Michigan; and David Geffen School of Medicine at UCLA, Los Angeles, California.

Data sets from the validation centers were retrospectively collected in compliance with local and national regulatory requirements and sent anonymized to the Paris Transplant

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Group. The study was approved by the Western Institutional Review Board (protocol number: 110989). Data were collected as part of an NIH-funded project (NIDDK R21DK122229). In France, the transplantation allocation system followed the rules of the French National Agency for Organ Procurement (Agence de la Bio-médecine). The European centers outside France followed the rules of the Eurotransplant allocation system (https://www.eurotransplant.org), and the US centers followed the rules of the US Organ Procurement and Transplantation System (https://unos.org/).

2.2. Kidney allograft failure risk evaluation

All patients included in the cohort underwent a posttransplant kidney allograft evaluation performed for clinical indication or as per protocol according to the centers' practices. Kidney allografts were evaluated at each allograft biopsy by retrospectively collecting clinical and biological data. A complete evaluation included the estimated glomerular filtration rate (eGFR) using the bedside Schwartz formula based on serum creatinine and height,[13] urine protein-to-creatinine ratio (UPCR), circulating anti-HLA DSA testing within 1 month of the biopsy, and a kidney allograft biopsy performed and interpreted according to the international Banff Classification of Renal Allograft Pathology.[14]

2.3. iBox score computation procedures

For each individual patient, we calculated a risk prediction score (iBox) according to the β regression coefficients estimated in the adult derivation cohort to obtain individual allograft survival probabilities up to 10 years after kidney allograft evaluation.[7] The iBox was calculated using 8 variables available at time of the kidney allograft evaluation: time from transplant to evaluation, functional data (eGFR and UPCR), immunological data (mean fluorescence intensity [MFI] of the immunodominant circulating anti-HLA DSA), and histological data, including microcirculation inflammation (glomerulitis [g] and peritubular capillaritis [ptc] Banff scores), interstitial inflammation and tubulitis (*i* and *t* Banff scores), transplant chronic glomerulopathy Banff score), and interstitial fibrosis and tubular atrophy Banff score. In addition, an abbreviated iBox score without the kidney allograft biopsy parameters was calculated using 4 variables available at time of the kidney allograft evaluation: time from transplant to evaluation, functional data (eGFR and UPCR), and immunological data (MFI of the immunodominant circulating anti-HLA DSA). In patients with multiple kidney allograft failure risk evaluations during the study period, we randomly selected one evaluation in the main analysis.

2.4. Sensitivity analyses

For sensitivity analysis, we assessed the predictions' performances at 10 years after evaluation by including only kidney allograft evaluation performed at 1 year posttransplant and including only kidney allograft evaluations performed before the age of 18 years.

2.5. Outcome measure

The outcome of interest was allograft failure, defined as return to dialysis or pre-emptive retransplantation. The duration of follow-up was from the patient's kidney allograft evaluation (starting point) to the date of allograft failure or the date of last follow-up. For

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

patients who died with a functioning allograft, allograft survival was censored at the time of death as a functional allograft. Patients who did not experience allograft failure were censored at last follow-up.

2.6. Statistical analysis

We describe continuous variables by using means and standard deviations or medians and interquartile ranges (IQR). We compared means and proportions between groups by using Student's t test, analysis of variance (Mann-Whitney test for MFI), or the $\chi 2$ test (or Fisher exact test when appropriate).

2.7. Prediction performance assessments

We assessed the accuracy of the prediction model based on its discrimination ability and calibration performance. We evaluated the discrimination ability (the ability to separate patients with different prognoses) of the final model by using Harrell's concordance index (C index).[9] We assessed calibration (the ability to provide unbiased survival predictions in groups of similar patients) based on a visual examination of the calibration plots by using the rms package in R. For sensitivity analysis, we assessed the predictions' performances, including only kidney allograft evaluation before the age of 18 years. We used SAS Enterprise and R, version 4.1.2, for all analyses and considered *P* values <.05 to be significant; all tests were 2 tailed.

3. Results

3.1. Multicenter international pediatric cohort characteristics

A total of 706 patients met the inclusion criteria, including 402 patients from Europe and 304 patients from the US. Overall, 62% were males, and the leading causes of kidney failure were congenital anomalies of the kidney and the urinary tract, glomerulonephritis, and genetic diseases, accounting for 37%, 20%, and 13%, respectively. Median age at transplantation was 12 years (IQR, 7–15 years); 491 (71%) patients received a transplant from a deceased donor, and the median age of the donors was 19 years (IQR, 14–34 years) (Table). A total of 188 (28%) patients were transplanted pre-emptively, and only 3% of the cohort had a history of prior kidney transplantation. Recipients from Europe were more likely to receive a deceased donor transplant (81% vs 60%, P <.001) from a pediatric donor (median age, 16 [12–27] years in Europe vs 29 [20–36] years in the US; P < .001), with fewer HLA mismatches (median HLA mismatch number in A/B/DR, 3 [3–4] in Europe vs 4 [3–5] in the US; P < .001) but a longer cold ischemia time (mean cold ischemia time, 14 [9–18] hours in Europe vs 9 [2–15] hours in the US; P < .001). Delayed graft function, defined as the need for dialysis within the first week posttransplant, was rare (12%), and no significant difference in prevalence was found between Europe and the United States (Table). The median time from transplant to kidney allograft evaluation was 9.1 months (IQR, 3.3–19.2 months). The distribution of posttransplant kidney allograft evaluation times is provided in Supplementary Figure S2. At the time of evaluation, the mean eGFR was 68.7 ± 28.1 mL/min/1.73 m^2, and the median UPCR was 0.1 [0.0–0.4] g/g, and 134 (16.9%) of the patients presented with circulating anti-HLA DSA. On the histologic evaluation, we found some microvascular inflammation (ptc ≥ 1 or g ≥ 1) in 156 (22%) biopsies, with a

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

severe microvascular inflammation (defined as g + ptc Banff score ≥ 3) in only 37 (5%). Interstitial inflammation and tubulitis were present in 170 (24%) and 189 (27%) of the biopsies, respectively. Considering chronic lesions, moderate-to-severe interstitial fibrosis/tubular atrophy was found in 108 (15%) of the biopsies, chronic vascular lesions (cv Banff score ≥ 1) in 171 (24%), and chronic glomerulopathy in 35 (5%) of the biopsies (Table). The median follow-up time from the date of transplantation to last patient follow-up was 64 months (IQR, 44–93 months).

3.2. Application and performance assessment of the iBox score in predicting allograft failure

Among the 706 kTx recipients included, 80 allografts failed during the follow-up. The iBox score showed good discrimination performance with a C-statistic of 0.825 (95% bootstrap percentile confidence interval [CI], 0.763–0.887), 0.811 (95% CI, 0.751–0.870), 0.810 (95% CI, 0.750–0.868), and 0.807 (95% CI, 0.749–0.864) at 3, 5, 7, and 10 years, respectively (Fig. 1). Visual inspection of the calibration plots showed good agreement between the iBox risk score predicted probabilities of allograft survival at 3, 5, 7, and 10 years after risk evaluation and actual kidney allograft survival (Fig. 2). In addition, the prediction performances of the abbreviated iBox score (using 4 parameters without kidney allograft biopsy parameters) were assessed. Although lower than that of the full iBox model, the discrimination of the abbreviated iBox remained good with a C-statistics of 0.818 (95% CI, 0.755–0.880), 0.803 (95% CI, 0.745–0.861), 0.801 (95% CI, 0.744–0.858), and 0.798 (95% CI, 0.739–0.857) at 3, 5, 7, and 10 years, respectively. Visual inspection of the calibration plots showed adequate agreement between the abbreviated iBox risk score predicted probabilities of allograft survival at 3, 5, 7, and 10 years after risk evaluation and the actual kidney allograft survival (Supplementary Fig. S3).

3.3. Sensitivity analysis

We performed a sensitivity analysis including only kidney allograft evaluation performed at 1-year posttransplant (n = 359, 50.8%) and found good discrimination with a C-statistic at 10 years after evaluation of 0.792 (95% CI, 0.709–0.875). The performance of the model was higher in indication biopsies and increased with time since transplantation (Supplementary Figs. S4 and S5). In addition, we also assessed the performance of the iBox score for kidney allograft evaluation before the age of 18 years (n = 651, 92.2%) and found consistent discrimination, with a C-statistic at 10 years after evaluation of 0.801 (95% CI, 0.736–0.865).

4. Discussion

In this study, we demonstrated that the risk prediction score for kidney allograft failure (iBox) is highly accurate in predicting allograft failure in pediatric kTx recipients and showed discriminations and calibrations similar to what was previously reported in various cohorts of adult kTx recipients around the world.[7] Our study shows the generalizability of the iBox and answers the urgent need for an accurate prediction tool for allograft survival in pediatric kTx recipients.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Very few previous studies focused specifically on developing or validating an allograft prediction score in pediatric transplant recipients, and none of these studies provided an external validation of these scores and therefore enough evidence to be used as a surrogate endpoint of allograft survival.[3] Moreover, these models were based on registry data that do not capture important posttransplant predictors of kidney allograft failure, such as DSA or allograft biopsy results. Among these studies, 2 scores derived from registry data (European (ERA-EDTA/ESPN) registry[5] and French (REIN) registry[6], respectively) reported good discrimination on the derivation cohort. When comparing the predictors included in these risk scores specifically derived from a pediatric population to the iBox, the main predictors not included in the iBox were the recipient age and predictors of the quality of the donor (donor type, donor age, and cold ischemia time). It is important to note that these predictors were included in the iBox derivation study but were not selected in the final model when histological markers likely to better account for allograft quality were included.[7] Indeed, various factors such as allograft hypoperfusion or medication adherence may mediate the association between age and allograft loss and are likely to be captured by other predictors included in the iBox although a potential independent effect of age cannot be totally ruled out based on our analysis.[15,16]

The validation of the iBox risk prediction score in pediatric recipients confirms the robustness of this score, which was already extensively validated in various cohorts of adult kTx recipients from Europe, the US, and South America and in different subpopulations (eg, based on the type of donor or the immunological risk) and clinical settings (eg, surveillance vs indication biopsy).[7] Although sensitivity analyses on subgroups are challenging to perform in children given the small number of kidney transplants performed each year, our findings on a diverse cohort of pediatric patients from Europe and the United States with various practices in terms of transplant immunosuppression regimens and in various clinical conditions (including both surveillance and indication biopsies) suggest that the iBox risk score may be used reliably in various clinical situations in the pediatric population (examples of use provided in Supplementary Fig. S6). As expected, the performance of the model increased with the time since transplantation and in case of indication biopsies. Indeed, pediatric patients usually receive optimal kidneys from young donors with low immunological risk, resulting in good function and histology early posttransplant with major injury to the allograft occurring after transplantation.

We also demonstrated overall good performance for the abbreviated iBox score that does not include biopsy data. However, we observed a lower calibration of this model, suggesting that this model will be less precise in predicting the exact risk of allograft loss. Nevertheless, this model will be interesting in clinical practice since it can be applied whenever blood and urine analyses are performed. This could be of particular interest in centers that do not perform surveillance biopsies or in patients at increased risk of complication following transplant biopsy.

The iBox risk prediction score was also initially validated in 6 RCTs, and we demonstrated that it was able to capture the effect of therapeutic interventions on long-term graft outcome.[8] This led to the qualification of the iBox as an efficacy endpoint in clinical trials.[11] RCTs among pediatric kTx recipients are particularly challenging given the low

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

prevalence of the condition, the small size of the transplant programs, and the good transplant survival rates, forcing investigators to extend follow-up and rely on endpoints other than graft survival. As an example, the CRADLE trial comparing everolimus with reduced tacrolimus and steroid withdrawal with standard of care involved 28 centers from 13 countries and included 106 patients.[17] Two coprimary outcomes were used: a composite efficacy failure endpoint comprising biopsy-proven acute rejection, graft failure, or death occurring between randomization and 12 months posttransplant, and eGFR at 12 months. Interestingly, the authors acknowledge as a limitation of their study that graft failure was rare and not different between the groups, so that biopsy-proven acute rejection was the main event analyzed in the trial. This illustrates the difficulty of performing clinical trials in the pediatric transplant population, explains why few trials focus on this population, and strongly advocates for the development and validation of reliable surrogate endpoints of allograft failure in pediatric kTx recipients. The validation of the iBox among pediatric kTx recipients opens new perspectives for clinical trials in children and may streamline the validation of new treatments in this population.

The main strength of our study is the constitution of a large, international, and unselected cohort of pediatric kTx recipients and the collection of granular histological and immunological data that are lacking from registry data. Moreover, by performing the validation of a score developed among adults, we both support the robustness of this score and provide a score that can be used longitudinally in the same patient from childhood to adulthood, which is of major interest among adolescents. A main limitation of our study is the low incidence of allograft failure in children, which precluded us from performing sensitivity analysis in various specific subgroups and the lack of validation cohorts from other parts of the world such as Asia, Africa, and South America. However, previous results in adults suggest that the iBox risk score may be used in various clinical situations and populations. Finally, despite the large cohort included in this study, the low incidence of allograft loss in pediatric patients prevented us from developing and externally validating a specific pediatric model. Moreover, we did not have all the parameters included in previous publications and could not directly compare our model with previously published pediatric models.

5. Conclusion

This study is the first to demonstrate the accuracy of the iBox system in predicting kidney allograft failure in children, with performances similar to those reported in adults. This supports its use to further improve patient monitoring and facilitate clinical trials by using the iBox as a surrogate outcome of allograft loss in pediatric trials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank people who contributed to the data collection: Ouardia Nekachtali (Robert Debré Hospital, APHP, Paris), Abderrahmane Adel Aidoud, Marion Rabant, Jean-Paul Duong Van Huyen

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Paris Transplant Group), Jean-Luc Taupin (Immunology department, Saint-Louis hospital, APHP, Paris), Margo Kamel, Mone Anzai (Emory University), Tin Duong Mattel Children's Hospital, UCLA), Lynn Minton (UTHSC/Le Bonheur Children's Hospital, Memphis, Tennessee), Lutz T. Weber (Pediatric Nephrology, Children's and Adolescents' Hospital, University Hospital of Cologne, Cologne, Germany), Atif Awan (Temple Street Children's University Hospital, Dublin, Ireland), Lars Pape (Pediatric Nephrology, Department of Pediatrics II, University Children's Hospital Essen, Essen, Germany), Jun Oh (Department of Pediatric Nephrology, University Children's Hospital, Hamburg, Germany), Kai Krupka, Burkhard Toenshoff (Department of Pediatrics I, University Children's Hospital, Heidelberg, Germany), Jens Christian König (Pediatric Nephrology, University Children's Hospital Münster, Münster, Germany), Andrea Carraro (Pediatric Nephrology, Dialysis and Transplantation Unit, Department of Woman's and Child's Health, University Hospital of Padova, Padua, Italy), Luca Dello Strologo (Pediatric Nephrology and Renal Transplant Unit, Bambino Gesù Children's Hospital-IRCCS, Rome, Italy), Matthias Zirngibl, Marcus Weitz (University Children's Hospital, Tübingen, Germany), Marcus Weitz (Pediatric Nephrology Department, University Children's Hospital Zurich, Zurich, Switzerland).

Funding

This study was supported by a grant from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK, R21DK122229). We gratefully acknowledge funding of the CERTAIN registry by a grant from the Dietmar Hopp Stiftung, the European Society for Paediatric Nephrology (ESPN), the German Society for Paediatric Nephrology (GPN), and grants from the pharmaceutical companies Astellas and Novartis.

Data availability statement

Data are available from the corresponding author at alexandreloupy@aphp.fr upon reasonable request for research purpose only.

Abbreviations:

CAKUT	congenital anomalies of the kidney and the urinary tract
cg	chronic glomerulopathy
DSA	donor-specific antigen
eGFR	estimated glomerular filtration rate
g	glomerulitis
HLA	human leukocyte antigen
IFTA	interstitial fibrosis and tubular atrophy
IQR	interquartile range
kTx	kidney transplant
MFI	mean fluorescence intensity
RCT	randomized controlled trial
SD	standard deviation
UPCR	urine protein-to-creatinine ratio

References

1. Chesnaye NC, Van Stralen KJ, Bonthuis M, et al. The association of donor and recipient age with graft survival in paediatric renal transplant recipients in a European Society for Paediatric

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Nephrology/European Renal Association-European Dialysis and Transplantation Association Registry study. Nephrol Dial Transplant. 2017;32(11):1949–1956. 10.1093/ndt/gfx261. [PubMed: 28992338]

2. Hart A, Lentine KL, Smith JM, et al. OPTN/SRTR 2019 annual data report: kidney. Am J Transplant. 2021;21(suppl 2):21–137. 10.1111/ajt.16502.
3. Kaboré R, Haller MC, Harambat J, Heinze G, Leffondré K. Risk prediction models for graft failure in kidney transplantation: a systematic review. Nephrol Dial Transplant. 2017;32(suppl 2):ii68–ii76. 10.1093/ndt/gfw405. [PubMed: 28206633]
4. Krikov S, Khan A, Baird BC, et al. Predicting kidney transplant survival using tree-based modeling. ASAIO J. 2007;53(5):592–600. 10.1097/MAT.0b013e318145b9f7. [PubMed: 17885333]
5. Lofaro D, Jager KJ, Abu-Hanna A, et al. Identification of subgroups by risk of graft failure after paediatric renal transplantation: application of survival tree models on the ESPN/ERA-EDTA Registry. Nephrol Dial Transplant. 2016;31(2):317–324. 10.1093/ndt/gfv313. [PubMed: 26320038]
6. Kaboré R, Ferrer L, Couchoud C, et al. Dynamic prediction models for graft failure in paediatric kidney transplantation. Nephrol Dial Transplant Off Publ Eur Dial Transpl Assoc Eur Ren Assoc. 2021;36(5):927–935. 10.1093/ndt/gfaa180.
7. Loupy A, Aubert O, Orandi BJ, et al. Prediction system for risk of allograft loss in patients receiving kidney transplants: international derivation and validation study. BMJ. 2019;366:l4923. 10.1136/bmj.l4923. [PubMed: 31530561]
8. Aubert O, Divard G, Pascual J, et al. Application of the iBox prognostication system as a surrogate endpoint in the TRANSFORM randomised controlled trial: proof-of-concept study. BMJ Open. 2021;11(10):e052138. 10.1136/bmjopen-2021-052138.
9. Naesens M, Budde K, Hilbrands L, et al. Surrogate endpoints for late kidney transplantation failure. Transpl Int. 2022;35:10136. 10.3389/ti.2022.10136. [PubMed: 35669974]
10. LOI determination letter. Food and Drug Administration (FDA). Accessed June 1, 2020. https://www.fda.gov/media/139300/download.
11. Qualification opinion for the iBox Scoring System as a secondary efficacy endpoint in clinical trials investigating novel immunosuppressive medicines in kidney transplant patients. European Medicines Agency (EMA). Accessed 19 December 2022. https://www.ema.europa.eu/en/documents/scientific-guideline/qualification-opinion-ibox-scoring-system-secondary-efficacy-endpoint-clinical-trials-investigating_en.pdf.
12. Plotnicki L, Kohl CD, Höcker B, et al. The CERTAIN registry: a novel, web-based registry and research platform for pediatric renal transplantation in Europe. Transplant Proc. 2013;45(4):1414–1417. 10.1016/j.transproceed.2013.01.007. [PubMed: 23726585]
13. Schwartz GJ, Muñoz A, Schneider MF, et al. New equations to estimate GFR in children with CKD. J Am Soc Nephrol. 2009;20(3):629–637. 10.1681/ASN.2008030287. [PubMed: 19158356]
14. 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(9):2318–2331. 10.1111/ajt.15898. [PubMed: 32463180]
15. Foster BJ, Dahhou M, Zhang X, Platt RW, Samuel SM, Hanley JA. Association between age and graft failure rates in young kidney transplant recipients. Transplantation. 2011;92(11):1237–1243. 10.1097/TP.0b013e31823411d7. [PubMed: 22124283]
16. Kaboré R, Couchoud C, Macher MA, et al. Age-dependent risk of graft failure in young kidney transplant recipients. Transplantation. 2017;101(6):1327–1335. 10.1097/TP.0000000000001372. [PubMed: 27482961]
17. Tönshoff B, Ettenger R, Strologo LD, et al. Early conversion of pediatric kidney transplant patients to everolimus with reduced tacrolimus and steroid elimination: results of a randomized trial. Am J Transplant. 2019;19(3):811–822. 10.1111/ajt.15081. [PubMed: 30125462]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

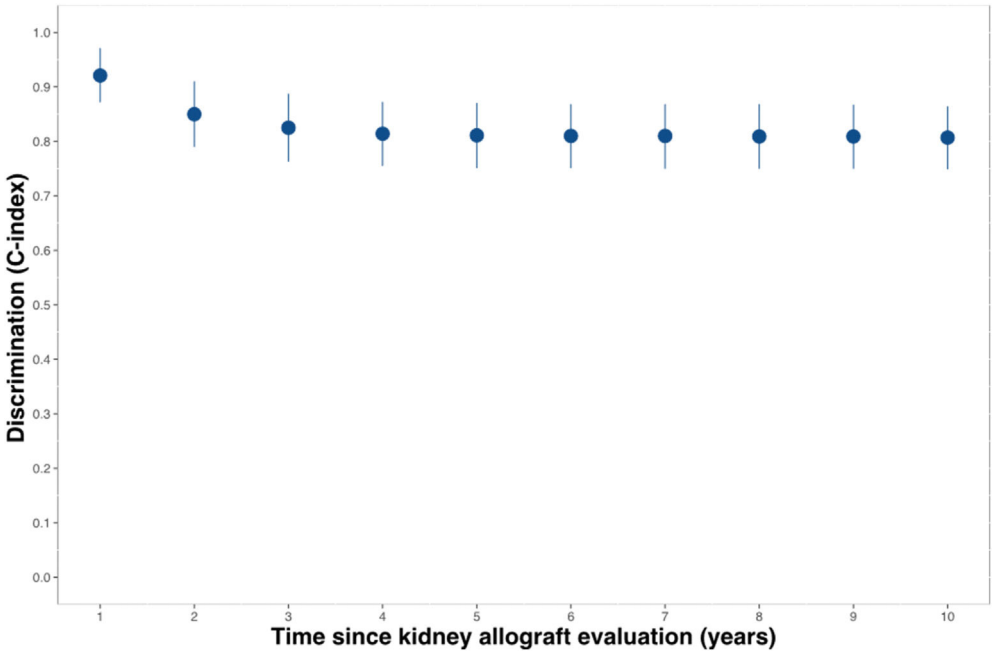

Figure 1.

Discrimination (Concordance Harrel C-index) of the integrative box risk prediction risk score to predict kidney allograft failure in the pediatric population from 1 to 10 years after kidney allograft evaluation. One-year C-index 0.921 (95% confidence interval [CI], 0.872–0.971), 2-year C-index 0.850 (95% CI, 0.790–0.910), 3-year C-index 0.825 (95% CI, 0.763–0.887), 4-year C-index 0.814 (95% CI, 0.755–0.872), 5-year C-index 0.811 (95% CI, 0.751–0.870), 6-year C-index 0.810 (95% CI, 0.751–0.868), 7-year C-index 0.810 (95% CI, 0.750–0.868), 8-year C-index 0.809 (95% CI, 0.750–0.868), 9-year C-index 0.809 (95% CI, 0.750–0.867), and 10-year C-index 0.807 (95% CI, 0.749–0.864). Each point with the range in blue represents the Concordance Harrel C-index with the 95% CI.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

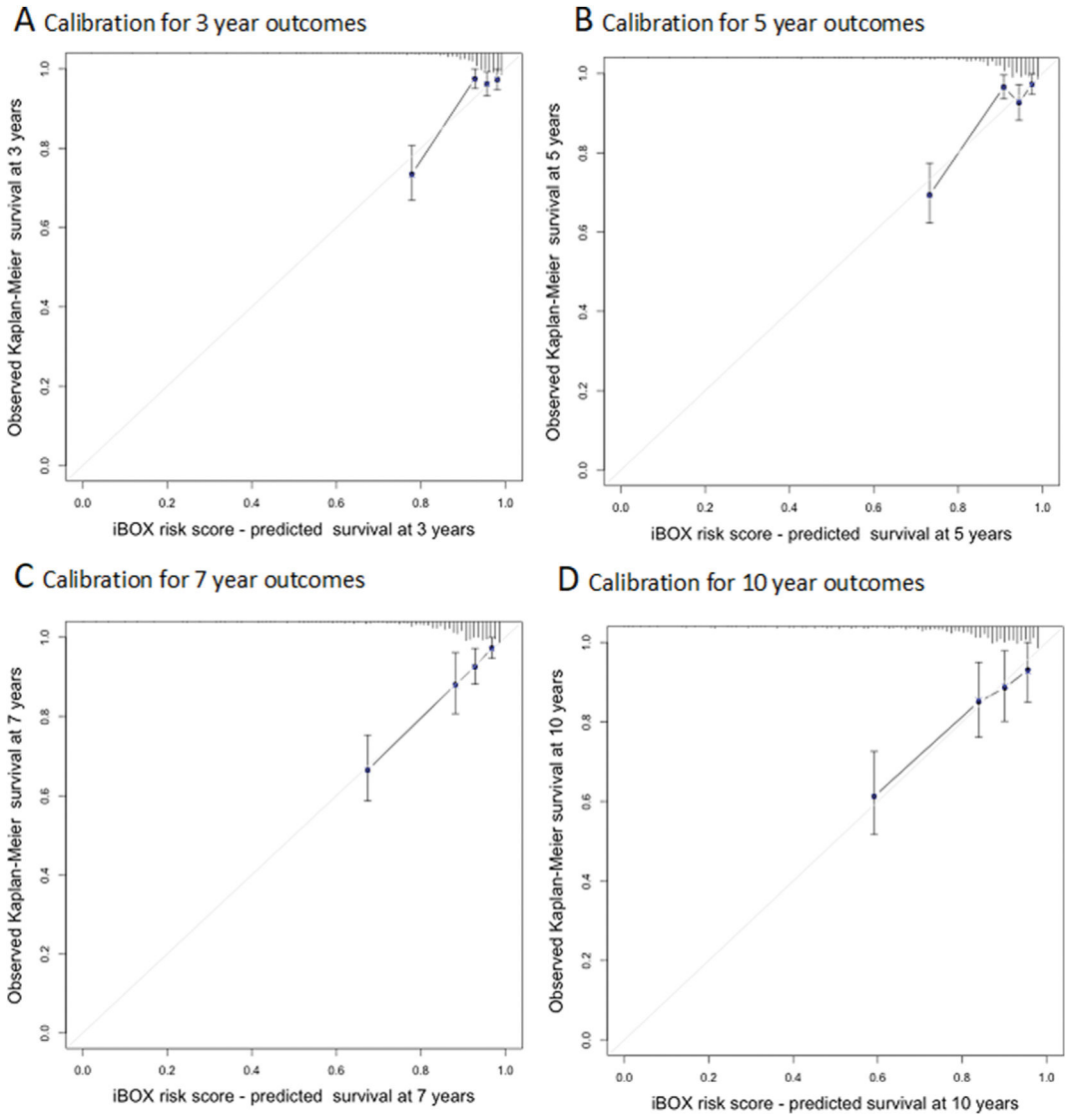

Figure 2.

Calibration plots at 3, 5, 7, and 10 years of the integrative box risk prediction system in the pediatric validation cohort after kidney allograft evaluation: (A) 3-year, (B) 5-year, (C) 7-year, and (D) 10-year predictions. Vertical axis is the observed proportion of grafts surviving at time of interest. Average predicted probability (predicted survival; x-axis) was plotted against Kaplan-Meier estimate (observed overall survival; y-axis). The grey line represents a perfectly calibrated model, and the dark line represents the integrative box risk prediction system.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Am J Transplant. Author manuscript; available in PMC 2024 October 01.

Table

Patients' and transplants' characteristic at the time of transplantation and at the time of kidney allograft evaluation.

Variable	Overall (n = 706)		European cohort (n = 402)	US cohort (n = 304)	*P* value
	N				
Recipient demographics					
Age at transplantation (y), median (IQR)	688	12 (7–15)	12 (7–15)	12 (6–16)	.96
Male sex, n (%)	688	426 (62)	228 (60)	198 (65)	.15
Kidney failure causes	706				<.001
CAKUT, n (%)		262 (37)	132 (33)	130 (43)	
Glomerulonephritis, n (%)		138 (19)	83 (21)	55 (18)	
Genetic diseases, n (%)		89 (13)	65 (16)	24 (8)	
Vascular, n (%)		34 (5)	23 (6)	11 (4)	
Other, n (%)		113 (16)	55 (14)	58 (19)	
Unknown, n (%)		70 (10)	44 (11)	26 (9)	
Transplant characteristics					
Donor age (y), median (IQR)	581	19 (14–34)	16 (12–27)	29 (20–36)	<.001
Donor type					
Deceased donor, n (%)	687	491 (71)	310 (81)	181 (60)	<.001
Pre-emptive transplant, n (%)	681	194 (28)	98 (26)	96 (32)	.05
Prior kidney transplant, n (%)	673	19 (3)	15 (4)	4(1)	.03
Cold ischemia time (h), mean (SD)	624	12 (4–18)	14 (9–18)	9 (2–15)	<.001
Delayed graft function, n (%)	514	63 (12)	41 (14)	22 (10)	.26
HLA-A/B/DR mismatch, median (IQR), number	628	4 (3–5)	3 (3–4)	4 (3–5)	<.001
Kidney allograft evaluation					
Time of evaluation (mo), median (IQR)	706	9.1 (3.3–19.2)	11.5 (3.4–21.1)	6.7 (3.1–13.8)	<.001
eGFR at the time of evaluation, mean (SD)	706	68.7 (28.1)	64.1 (24.9)	74.8 (30.8)	<.001
UPCR at time of evaluation (g/g), median (IQR)	706	0.01 (0.0–0.04)	0.02 (0–0.07)	0 (0–0.02)	<.001
DSA positive at the time of evaluation, n (%)	706	134 (16.9)	100 (24.9)	34 (11.2)	<.001
MFI at time of evaluation, n (%)					<.001
500–1000		25 (18.7)	25 (25.0)	0	
1000–3000		42 (31.3)	30 (30.0)	12 (35.3)	

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Variable	Overall (n = 706)	European cohort (n = 402)	US cohort (n = 304)	*P* value
	N			
>3000	67 (50.0)	45 (45.0)	22 (64.7)	
Microcirculation inflammation (g + ptc Banff scores), n (%)	706			.003
0–2	669 (94.8)	372 (92.5)	297 (97.7)	
3–4	30 (4.2)	23 (5.7)	7 (2.3)	
5–6	7 (1.0)	7 (1.7)	0	
Interstitial inflammation and tubulitis: (i + t Banff scores), n (%)	706			.432
0–2	624 (88.4)	352 (87.6)	272 (89.5)	
≥3	82 (11.6)	50 (12.4)	32 (10.5)	
Transplant chronic glomerulopathy (cg Banff scores)	706			.034
0	671 (95.0)	376 (93.5)	295 (97.0)	
≥1	35 (5.0)	26 (6.5)	9 (3.0)	
IFTA Banff score	706			.004
0–1	598 (84.7)	332 (82.6)	266 (87.5)	
2	76 (10.8)	43 (10.7)	33 (10.9)	
3	32 (4.5)	27 (6.7)	5 (1.6)	

CAKUT, congenital anomalies of the kidney and the urinary tract; cg, chronic glomerulopathy; DSA, donor-specific antigen; eGFR, estimated glomerular filtration rate; g, glomerulitis; HLA, human leukocyte antigen; i, interstitial inflammation; IFTA, interstitial fibrosis and tubular atrophy; IQR, interquartile range; MFI, mean fluorescence intensity; ptc, peritubular capillaritis; SD, standard deviation; t, tubulitis; UPCR, urine protein-to-creatinine ratio; US, United States.

Am J Transplant. Author manuscript; available in PMC 2024 October 01.