

OPEN

Pretransplant Donor-specific IFN γ ELISPOT as a Predictor of Graft Rejection: A Diagnostic Test Accuracy Meta-analysis

Nuria Montero, MD, PhD,¹ Samira Farouk, MD,² Ilaria Gandolfini, MD,³ Elena Crespo, PhD,⁴ Marta Jarque, MS,⁴ Maria Meneghini, MD,⁴ Alba Torija, MS,⁴ Umberto Maggiore, MD, PhD,³ Paolo Cravedi, MD, PhD,² and Oriol Bestard, MD, PhD^{1,4}

Background. Pretransplant interferon- γ enzyme-linked immunospot (IFN- γ ELISPOT) has been proposed as a tool to quantify alloreactive memory T cells and estimate the risk of acute rejection (AR) after kidney transplantation, but studies have been inconclusive so far. We performed a meta-analysis to evaluate the association between pretransplant IFN- γ ELISPOT and AR and assess its predictive accuracy at the individual level. **Methods.** We estimated the pooled summary of odds ratio for AR and the joined sensitivity and specificity for predicting AR using random-effects and hierarchical summary receiver-operating characteristic models. We used meta-regression models with the Monte Carlo permutation method to adjust for multiple tests to explain sensitivity and specificity heterogeneity across studies. The meta-analytic estimates of sensitivity and specificity were used to calculate positive and negative predictive values across studies. **Results.** The analysis included 12 studies and 1181 patients. IFN- γ ELISPOT was significantly associated with increased AR risk (odds ratio: 3.29; 95% confidence interval (CI), 2.34-4.60); hierarchical summary receiver operating characteristic jointly estimated sensitivity and specificity values were 64.9% (95% CI, 53.7%-74.6%) and 65.8% (95% CI, 57.4%-73.5%), respectively, with moderate heterogeneity across studies. After adjusting for multiple testing, meta-regression models showed that thymoglobulin induction, recipient black ethnicity, living versus deceased donors, and geographical location did not affect sensitivity or specificity. Because of the varying AR incidence of the studies, positive and negative predictive values ranged between 16%–60% and 70%–95%, respectively. **Conclusions.** Pretransplant IFN- γ ELISPOT is significantly associated with increased risk of AR but provides suboptimal predictive ability at an individual level. Prospective randomized clinical trials are warranted.

(*Transplantation Direct* 2019;5:e451; doi: 10.1097/TXD.0000000000000886. Published online 25 April, 2019.)

During the last decades, the refinement of prekidney transplant risk immune monitoring and evaluation of the humoral arm of adaptive immunity has led to more effective kidney allocation and a dramatic reduction

of posttransplant acute antibody-mediated rejection.¹⁻³ However, T cell-mediated rejection (TCMR) unpredictably occurs, as no immune-risk stratification of cellular sensitization before transplantation is available in current clinical

Received 30 January 2019.

Accepted 18 February 2019.

¹ Nephrology Department, Kidney transplant Unit, Bellvitge University Hospital, Barcelona, Spain.

² Department of Medicine, Translational Transplant Research Center, Icahn School of Medicine at Mount Sinai, New York, NY.

³ Trapianti rene pancreas (UO Nefrologia), Dipartimento di Medicina e Chirurgia, Università di Parma, Parma, Italy.

⁴ Experimental Nephrology Laboratory, IDIBELL, Barcelona, Spain.

N.M., S.F., and I.G. are co-first authors.

P.C. and O.B. are co-senior authors.

N.M. drafted the protocol, planned the study, selected and assessed the studies, performed the analysis, and prepared the article. S.F. selected and assessed the studies, extracted data, and prepared the article. I.G. selected and assessed the studies, extracted data, and prepared the article. E.C., M.J., M.M., and A.T. assisted in article review. U.M. drafted the protocol, planned the study, performed the analysis, and prepared the article. P.C. and O.B. drafted the protocol, planned the study, assisted in the data analysis, prepared the article, and reviewed the final manuscript. All authors read and approved the final manuscript.

The authors declare no conflicts of interest.

This work was partially supported by 2 Spanish grants from "Instituto de Salud

Carlos III" (INT15/00112; PI16/01321) and 2 European Commission grants from the Biomarker-Driven Immunosuppression Minimization (BIODRIM) Consortium (FP7/2007–2013) and the Horizon 2020 EU research and innovation program (EU-TRAIN; 18CEE002).

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: Paolo Cravedi, MD, PhD, Icahn School of Medicine at Mount Sinai, 1 Levy Place, 10029 New York, NY.

(paolo.cravedi@mssm.edu) or Oriol Bestard, MD, PhD, Nephrology Department, Kidney Transplant Unit, Bellvitge University Hospital, Feixa Llarga s/n 08907. Barcelona, Spain. (obestard@bellvitgehospital.cat).

Copyright © 2019 The Author(s). *Transplantation Direct*. Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

ISSN: 2373-8731

DOI: 10.1097/TXD.0000000000000886

practice. Indeed, immunosuppressive protocols after kidney transplantation are generally adjusted on empirical bases and on a functional or histological evaluation of the allograft and/or signs of drug toxicity or infection.⁴ As a result, there are patients who are likely to receive too much or too little immunosuppression—exposing them to higher rates of infection, malignancy and drug toxicity, or conversely, to increased risk of acute and chronic graft rejection. Therefore, developing reliable biomarkers of antidonor T-cell alloimmune reactivity is crucial to individualize immunosuppressive therapy ultimately aimed at reducing allograft rejection risk, while minimizing adverse effects associated with over-immunosuppression.⁵

The interferon- γ enzyme-linked immunospot (IFN- γ ELISPOT) is an immune assay that has been developed and standardized in the field of transplantation as a way to functionally measure the number of circulating memory T cells with donor antigen reactivity.^{6,7} While a number of single-center studies have reported a close association between positive pretransplant IFN- γ ELISPOT and increased risk of acute rejection (AR), particularly TCMR, and poor graft function at 6 and 12 months after transplantation,⁸⁻¹⁰ other studies have failed to identify such associations.¹¹ Due to the high variability in patient characteristics and treatments across studies, we hypothesized that the IFN- γ ELISPOT performs differently in various patient populations and settings. Moreover, most published studies lack an acceptable balance regarding the predictive accuracy, which is a key for a biomarker to be used in the clinical practice.

To overcome these issues, we performed a meta-analysis of aggregate data from all published studies (including almost 1200 kidney transplant patients) evaluating the predictive performance of the pretransplant donor-specific IFN- γ ELISPOT as the most promising immune biomarker assessing the risk of TCMR in kidney transplant recipients. Our primary objective was to provide precise estimates of the predictive value of the pretransplant IFN- γ ELISPOT for AR at the individual level, taking into account all potential sources of heterogeneity across all the studies that may ultimately influence its predictive accuracy.

MATERIALS AND METHODS

Literature Search

We performed a systematic literature search using the Cochrane Central Library, MEDLINE, and EMBASE databases up to October 2016 using a predefined algorithm (Table S1, SDC, <http://links.lww.com/TXD/A207>). Abstracts from proceedings of major conferences, International Clinical Trials Registry Platform Search Portal, and ClinicalTrials.gov were also searched. References included in pertinent systematic reviews were then screened. Studies were deemed eligible if they measured prekidney transplant donor-specific IFN- γ ELISPOT in patients who received a living- or deceased-donor kidney transplant alone. We excluded the studies if immunosuppression was modified on the basis of ELISPOT results.¹² During screening, only articles written in English were included to prevent any misinterpretation of data. Study outcomes were not included as a part of the eligibility criteria.

All references were screened by 2 independent reviewers (S.F. and I.G.). If any discrepancies occurred, 2 additional

reviewers were consulted (N.M. and O.B.). The study reviewers also examined reference lists of clinical practice guidelines, review articles, and relevant studies to identify missing articles and sent e-mails to investigators known to be involved in previous studies seeking information about unpublished or incomplete trials.

The protocol of this systematic review is published in the PROSPERO register (#CRD42018116382).

Data Extraction, Outcomes, and Quality Assessment

Data extraction was independently performed by 3 reviewers (N.M., S.F., and I.G.) using standard data extraction forms. Reviewers were not blinded to authors, institutions, or article journals. If any discrepancies occurred, a consensus was reached after consulting a fourth investigator (O.B.).

Data extraction included the following characteristics of the studies: author's name, year of publication, number of included patients, duration of follow-up, donor and recipient demographics, and clinical characteristics. Outcomes were identified in the studies' text or tables. Data regarding death and graft loss were extracted from survival curves when necessary. The primary outcome was the incidence of AR (both clinically suspected and biopsy proven). Secondary outcomes were graft function (defined by estimated glomerular filtration rate [eGFR] and serum creatinine) and patient and graft survival.

Risk of Bias Assessment/Quality Assessment

We used Hayden criteria to estimate the bias of the included studies.¹³

This tool evaluates 6 domains as follows:

1. Study participation: the study sample adequately represents the population of interest.
2. Study attrition: the study data available (ie, participants not lost to follow-up) adequately represent the study sample.
3. Prognostic factor (PF) measurement: the PF is measured in a similar way for all participants.
4. Outcome measurement: the outcome of interest is measured in a similar way for all participants.
5. Study confounding: important potential confounding factors are appropriately accounted for.
6. Statistical analysis and reporting: the statistical analysis is appropriate, and all primary outcomes are reported.

Data Synthesis and Analysis

We presented the results from all studies using the original manufacturer scale for test interpretation. The IFN- γ ELISPOT value cutoff to define positivity in the majority of studies was 25 spots/ 3×10^5 peripheral blood mononuclear cells (PBMCs),^{9,11,12,14-18} though in 2 studies the cutoff was defined as 12 spots/ 3×10^5 PBMCs.^{19,20} In 1 study,¹⁸ we evaluated a cutoff of 25 spots/ 3×10^5 PBMCs to homogenize all study data, as the cut-off was not explicitly defined.

We followed the Meta-analyses of Observational Studies in Epidemiology guidelines to report this systematic review and used Stata SE Release 15 (StataCorp, College Station, TX) for all analyses.²¹

We evaluated the association between pretransplant donor-specific IFN- γ ELISPOT and AR by calculating the odds ratio (OR) via random-effects meta-analysis. To overcome the presence of 0 cells in some of the 2×2 contingency tables, we added a small constant $\epsilon = 0.5$ to all cells.²²

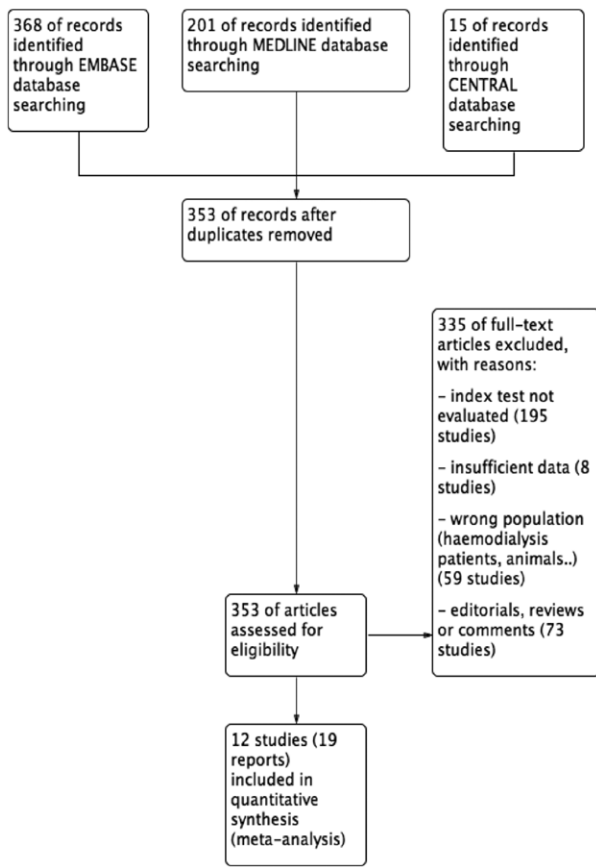


FIGURE 1. Flow of the studies reviewed.

To examine the predictive accuracy of pretransplant IFN- γ ELISPOT at the individual level, we performed a random-effects meta-analysis on sensitivity and specificity. Sensitivity represents the proportion of IFN- γ ELISPOT-positive patients who eventually developed AR, whereas specificity is the proportion of IFN- γ ELISPOT-negative patients who were AR free. To provide a joint pooled estimate of sensitivity and specificity which accounts for the correlation between the 2 parameters, we fitted a hierarchical summary receiver operating characteristic (ROC) model using the user-written Stata command *metandi*.^{23,24} Using the command *metandplot*, we also plotted the model-fitted results, a summary ROC curve, a 95% confidence region, and a 95% prediction region (ie, confidence region for a forecast in a future study of the summary operating point of sensitivity and specificity).^{23,24} Besides calculating the meta-analytic estimates of joint sensitivity and specificity, we also calculated the meta-analytic positive and negative likelihood ratios. These ratios can be interpreted as follows: a completely nonpredictive assay is reflected by a positive and negative likelihood ratio of 1; a slightly, moderately, and highly predictive assay is reflected by a positive likelihood ratio of above 2, 5, and 10 and by a negative likelihood ratio of below 0.5, 0.2, and 0.1, respectively.²⁵

In an attempt to explain the reasons for heterogeneity across study findings concerning sensitivity and specificity and to identify patient categories in which pretransplant IFN- γ ELISPOT could be most useful, we additionally fitted metaregression models using the user-written Stata *metareg* command.²⁶ These models were designed to examine to what extent sensitivity and specificity of IFN- γ ELSPOT could be affected by

TABLE 1. Study characteristics

References (country)	Study design	N	Follow-up, mo	Recipient age, mean (SD)	Recipient sex, % F	Recipient ethnicity, % white	Recipient ethnicity, % AA	Cadaveric, %	DSA %	+ ELISPOT spots/PBMC	T cell-depleting induction, %	CNI maintenance, %
Augustine et al ¹⁴ (United States)	Single center, observational	37	12	47	37.8	0	100	100	NR	>25/3×10 ⁶	16.1	100
Augustine et al ¹⁵ (United States)	Single center, observational	100	12	48 (13)	37	NR	53	60	NR	>25/3×10 ⁶	8.6	100
Augustine et al ¹⁶ (United States)	Single center, observational	130	12	46.7	39.5	NR	53	61.5	NR	>25/3×10 ⁶	6.2	100
Crespo et al ¹⁰ (Spain)	Retrospective, single center	90 (T), 67 (V)	21 (mean, T)	54.8 (T), 55.2 (V)	31.2	93.6	NR	58	5.7	>25/3×10 ⁶	22.9	88.5
Crespo et al ¹⁷ (Spain)	Prospective, observational	101 (V)	24	56.3	34.6	94.1	NR	71.3	0	>25/3×10 ⁶	18.1	99
Gandolfini et al ⁸¹ (United States)	Retrospective, cohort study	168	45	56.2 (13.4)	35.7	94.6	NR	79.2	NR	>25/3×10 ⁶	24.4	100
Hricik et al ³⁰ (United States)	Retrospective, single center	118	12	47.5 (11.8)	37	NR	52	100	NR	>25/3×10 ⁶	23.7	NR
Hricik et al ¹¹ (United States)	Prospective, observational	176	12	40 (12.93)	51.7	NR	31.3	31.8	NR	>25/3×10 ⁶	48.2	84
Hricik et al ²⁹ (United States)	Randomized control trial	12	24	45.2 (11.31)	52.4	81	NR	0	NR	>25/3×10 ⁶	100	CNI withdrawal RCT
Kim et al ²⁰ (Korea)	Observational, single center	45	6	41.2	28.9	NR	NR	0	NR	>12/2×10 ⁶	NR	100
Kim et al ¹⁹ (Korea)	Retrospective, cohort study	154	36	41.1	39	NR	NR	0	NR	>12/2×10 ⁶	0	70.8
Nickel et al ¹⁸ (United States)	Retrospective, single center	52	12	49.9 (11)	36.5	NR	NR	84.6	NR	NR	3.8	100

AA, African American; CNI, calcineurin inhibitor; F, female; NR, not reported; PBMC, peripheral blood mononuclear cell; RCT, randomized controlled trial; T, training set; V, validation set.

TABLE 2.**Risk of bias of the included studies based on the Hayden criteria**

Bias	Nickel et al ¹⁸	Augustine et al ¹⁴	Augustine et al ¹⁵	Kim et al ²⁰	Augustine et al ¹⁶	Kim et al ¹⁹	Bestard et al ⁷	Hricik et al ³⁰	Crespo et al ¹⁰	Hricik et al ^{11,29}	Hricik et al ^{11,29}	Crespo et al ¹⁷
Study participation												
Adequate participation in the study by eligible persons	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Description of the source population or population of interest	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Description of the baseline study sample	Low	Mod	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Adequate description of the sampling frame and recruitment	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Adequate description of the period and place of recruitment	Low	Low	Low	High	Low	Mod	Low	Low	Low	Low	Low	Low
Adequate response rate for study participants	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Study attrition												
Description of attempts to collect information on participants who dropped out	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Reasons for loss to follow-up are provided	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Adequate description of participants lost to follow-up	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
No important differences between participants who completed the study and those who did not	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Mod	Low
Prognostic factor measurement												
A clear definition or description of the PF is provided	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Method of PF measurement is adequately valid and reliable	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Continuous variables are reported or appropriate cut points are used	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
The method and setting of measurement of PF is the same for all study participants	Low	Low	Low	Low	Mod	Low	Mod	Low	Low	Mod	Low	Low
Adequate proportion of the study sample has complete data for the PF	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low

Continued next page

TABLE 2. (Continued)

Bias	Nickel et al ¹⁸	Augustine et al ¹⁴	Augustine et al ¹⁵	Kim et al ²⁰	Augustine et al ¹⁶	Kim et al ¹⁹	Bestard et al ⁷	Hricik et al ²⁰	Crespo et al ¹⁰	Hricik et al ^{11,29}	Hricik et al ^{11,29}	Crespo et al ¹⁷
Outcome measurement												
	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low
Appropriate methods of imputation are used for missing PF data												
A clear definition of the outcome is provided	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Method of outcome measurement used is adequately valid and reliable	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
The method and setting of outcome measurement is the same for all study participants	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low
Study confounding												
All important confounders are measured	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Clear definitions of the important confounders measured are provided	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Measurement of all important confounders is adequately valid and reliable	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
The method and setting of the same for all study participants	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Appropriate methods are used if imputation is used for missing confounder data	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Important potential confounders are accounted for in the study design	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Important potential confounders are accounted for in the analysis	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Sufficient presentation of data to assess the adequacy of the analytic strategy	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Strategy for model building is appropriate and is based on a conceptual framework or model	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
The selected statistical model is adequate for the design of the study	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
There is no selective reporting of confounding measurement are of results	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low

^aInterpretation of domain-level and overall risk of bias judgments in Hayden criteria. Mod, moderate; PF, prognostic factor.

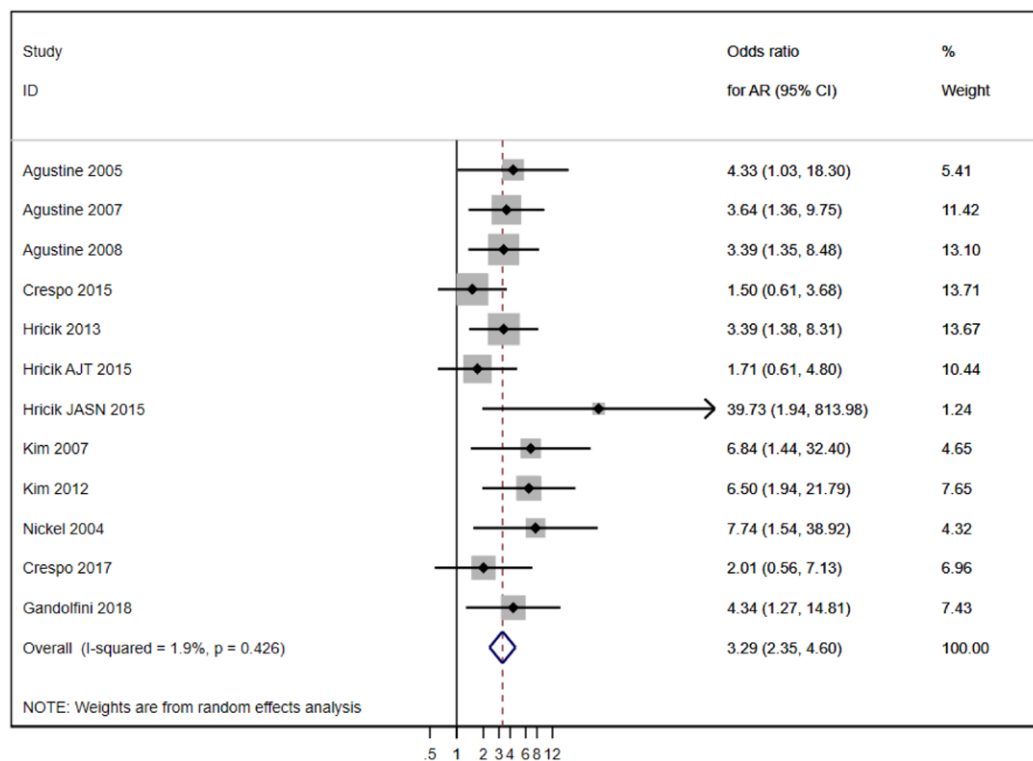


FIGURE 2. Forest plot of analysis acute rejection (AR).

the type of induction therapy used (independent variable = % of thymoglobulin use), by ethnicity (% of African American recipients), by living- versus deceased-donor kidney transplantation (% of living donors), by era of study publication (year as continuous variable), and by geographical location of the study (indicator variable for North America, Asia, or Europe). To avoid inflating the rate of false-positive findings because of multiple testing, we referred to Monte Carlo permutation tests rather than the default nominal P values when declaring statistical significance (ie, 2-tailed, $P < 0.05$) using the methodology proposed by Higgins and Thompson.²⁷ However, in the meta-regression plots, we reported default nominal P values calculated by the default Knapp-Hartung variance estimator and associated t tests.²⁸ For the purpose of meta-regression analyses, sensitivity and specificity were logit transformed. Sensitivity analyses were performed after the exclusion of the study published by our group.¹² In this study, the type of immunosuppression was given according to the pretransplant IFN- γ ELISPOT. To see whether the inclusion of a study in which some patients underwent complete tacrolimus withdrawal might have accounted for inconsistencies between study findings, the analysis was repeated without this study (Figure S1, SDC, <http://links.lww.com/TXD/A207>).

Heterogeneity was analyzed using a chi-squared test on $N-1$ degrees of freedom, with an alpha of 0.05 used for statistical significance and the I^2 test.²⁷ The I^2 statistic describes the percentage of total variation across studies that is due to heterogeneity rather than chance; I^2 values of 25%, 50%, and 75% are indicative of low, moderate, and extreme heterogeneity, respectively.²⁷ The pooled meta-analytic estimates should be interpreted with caution whenever I^2 is $>50\%$. An $I^2 >50\%$ indicates that most of the observed differences between the study findings on the relation between IFN- γ ELISPOT and

AR may be due to true underlying inconsistencies between study results rather than random variation.

Finally, the meta-analytic estimates of joined sensitivity and specificity were used to calculate positive predictive value (PPV) and negative predictive value (NPV) in each study population from the selected studies. Because PPV and NPV are functions not only of IFN- γ ELISPOT sensitivity and specificity but also of the population-specific incidence of AR, we did not report them as fixed numbers but rather as numbers varying with pretest risk of AR.

Similar analyses were performed to examine the secondary outcomes including the association between pretransplant IFN- γ ELISPOT and graft function, graft survival, and patient survival. We had planned to similarly analyze de novo DSA, but reported data in the studies were insufficient.

RESULTS

Literature Search Results and Study Characteristics

A total of 584 potentially relevant citations were identified (EMBASE: 368; CENTRAL: 15; MEDLINE: 201). After the review of the titles and abstracts and the removal of duplicates, a total of 353 potentially eligible articles were identified. We excluded 335 studies for the following reasons: index test not evaluated (195 studies); insufficient data (lack of sufficient data to derive 2×2 tables or assess test performance in participants) or duplicate data (8 studies); different target population (hemodialysis patients, animal models) (59 studies); or studies using pretransplant IFN- γ ELISPOT for treatment decision purposes (1 study).¹² The same procedure was followed for editorials, reviews, or comments (73 studies). We ultimately included 12 studies in our analyses (19 reports) with 1181 patients.^{10,11,14-20,29-38} Figure 1 shows the flow chart

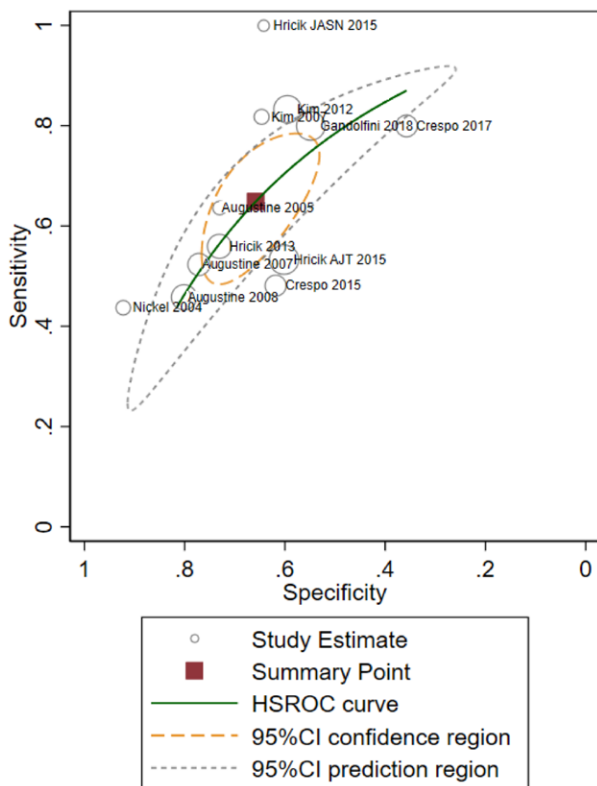


FIGURE 3. HSROC plot of analysis of acute rejection, namely joined meta-analysis of sensitivity and specificity and associated summary ROC curve concerning the predictive power of IFN γ ELISPOT for acute rejection, based on the fitted hierarchical logistic regression model. The graph shows the pooled summaries, together with circles showing the individual study estimates; the solid line represents the estimate of the summary ROC curve, and the square symbol represents a summary value for sensitivity and specificity of 64.9% (95% CI, 53.7-74.6) and 65.8% (95% CI: 57.4-73.5), respectively. The dotted green line represents the 95% confidence region for the summary operating point (ie, every sensitivity/specificity value lying within the region delimited by the green line must be regarded as not significantly different from our pooled estimate, at 5% significance level). The dotted orange line represents the 95% prediction region (confidence region for a forecast of the true sensitivity and specificity in a future study). The shape of the prediction region is dependent on the assumption of a bivariate normal distribution for the random effects included in the statistical model. It should therefore not be overinterpreted; it is intended to give a visual representation of the extent of between-study heterogeneity. CI, confidence interval; HSROC, hierarchical summary receiver-operating characteristic; IFN- γ , interferon- γ enzyme-linked immunospot; ROC, receiver operating characteristic.

for study selection. All studies reported pretransplant donor-specific IFN- γ ELISPOT data and included a total of 1238 kidney transplant patients. The characteristics of the included studies are presented in **Table 1**.

Risk of Bias

To evaluate the risk of bias, we followed the Hayden criteria.¹³ All the bias domains for all the studies are presented in **Table 2**.

In general, there was an overall low risk of bias in all the domains. The relationship between the PF and outcome may have differed between participants and eligible nonparticipants in 2 studies. In the study by Augustine,¹⁴ there was no information about the sampling frame and recruitment. The

authors of the study by Kim¹⁹ did not specify the place of recruitment. We found a high risk of bias in 1 study, as the authors did not give any information about the period and place of recruitment²⁰ and in another²⁹ because there was an early interruption with high incidence of rejections in the tacrolimus withdrawal arm. This reason made us consider also a moderate risk of bias due to presence of differences between participants who completed the study and those who did not.

The methods and setting of measurement of the PF in our meta-analysis (pretransplant IFN- γ ELISPOT) were not the same for all study participants, and that brought a moderate risk of bias in 3 studies.^{11,12,16} The method of imputation used for missing PF data was not correct and gave another study a high risk of bias.¹¹ The outcome measurement was of low risk of bias for all studies except for 1¹¹ that presents a high risk of bias in the method and setting of outcome measurement, which was not the same for all study participants.

Acute Rejection

Out of the 1181 included kidney transplant recipients, 209 (18%) developed AR, with an incidence ranging between 8% and 44% in each study. Pretransplant donor-specific IFN- γ ELISPOT was positive in 512 (43%) of the patients and ranged between 25% and 88% in each study.

Positive pretransplant IFN- γ ELISPOT was associated with a significantly high risk of AR (odds ratio 3.29; 95% confidence interval (CI), 2.34-4.60) (**Figure 2** and **Figure S1, SDC**, <http://links.lww.com/TXD/A207>). Pooled sensitivity and specificity for AR, jointly estimated via the hierarchical summary ROC model, were 64.9% (95% CI, 53.7%-74.6%) and 65.8% (95% CI, 57.3%-73.5%). The joint sensitivity and specificity estimates resulted in a positive likelihood ratio and a negative likelihood ratio of 1.90 (95% CI, 1.58-2.28) and 0.53 (95% CI, 0.42-0.68), respectively. Because the values of positive likelihood ratio below 2.0 and negative likelihood ratio above 0.5 are generally regarded as reflecting limited prognostic ability, our findings indicate that the added value of pretransplant IFN- γ ELISPOT to predict AR at the individual level is at the edge of being significant.³⁹ Notably, results of the joined meta-analysis of sensitivity and specificity and the associated summary ROC curve showed a certain degree of heterogeneity across studies (**Figure 3** and **Figure S2, SDC**, <http://links.lww.com/TXD/A207>).

To test whether differences across studies concerning sensitivity and specificity could be explained by differences in the characteristics of the study populations such as the type of induction therapy, ethnicity, and geographical area, and also to identify patient categories in which pretransplant IFN- γ ELISPOT could be most useful, we performed metaregression analyses (**Figure 4** and **Figures S3 and S4, SDC**, <http://links.lww.com/TXD/A207>). After adjusting for multiple testing to prevent spurious findings, metaregression models did not show evidence that thymoglobulin induction, recipient black ethnicity, living versus deceased donors, or geographical location affected sensitivity or specificity, but publication year was the only variable with a significant inverse association with specificity ($P = 0.021$).

We used the joint meta-analytic estimates of sensitivity and specificity to calculate PPV and NPV in each study population from the 12 selected studies. As a consequence of the fact that the study populations differed greatly with regard to the incidence of AR, PPV was rather low, ranging between 16%

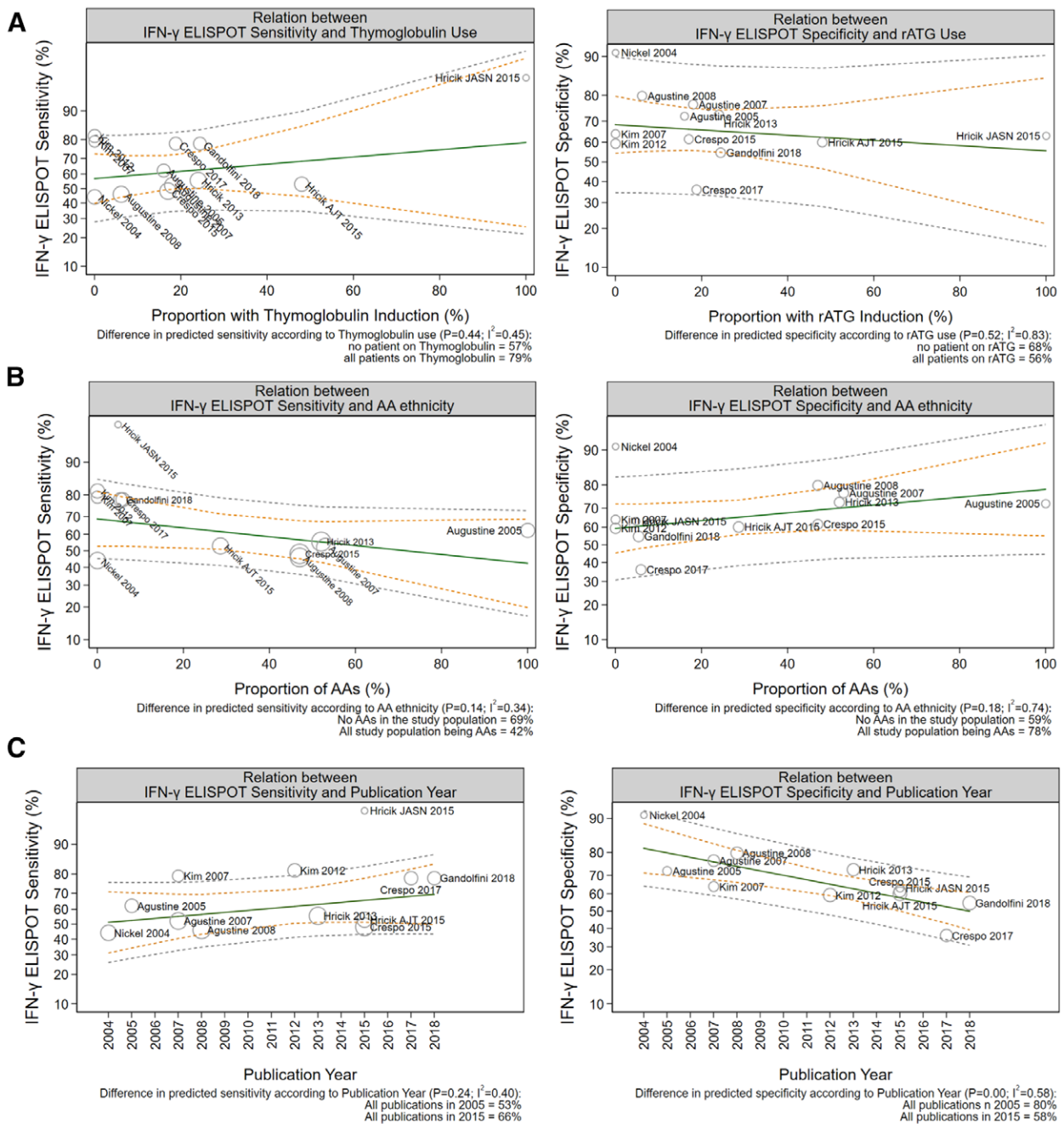


FIGURE 4. Relationship between sensitivity or specificity and the prevalence of thymoglobulin use (A), the prevalence of African Americans (AAs) (B), the publication year (C) of each study, estimated by random-effects meta-analysis (ie, metaregression). The size of the circles is proportional to the weight of each study, which is inversely proportional to the study variance. The variance is calculated as the sum of the within- and the between-study variances; therefore, it is not only dependent on the size of the study population. Sensitivity and specificity values on y-axis are obtained from backtransformation of the logit values; therefore, the scale of the y-axis is not linear. At the bottom of the graph are reported predictions based on the fitted model, for example, what would have happened to sensitivity or specificity if the prevalence of thymoglobulin use in the study population had been 0% or 100% (A); if the prevalence of AA had been 0% or 100% (B); and if the publication year of the studies had been 2005 for all studies or 2015 for all studies (C). Each figure reports the nominal P value using the default Knapp-Hartung variance estimator, testing that the regression slope is significantly different from zero. However, statistical inference mentioned in the text was not based on those nominal P values but rather on adjustment for multiple testing using Monte Carlo permutation test based on the methodology proposed by Higgins and Thompson. The I^2 represents the proportion of the variance between studies, which is due to unmeasured heterogeneity between studies rather than chance or difference in the variable on the x-axis of the plot. IFN- γ ELISPT, interferon- γ enzyme-linked immunospot.

and 60%, whereas NPV was significantly higher and ranged between 70% and 95% (Figure 5 and Table 3), respectively. In studies where the risk of AR was 16% or lower—which is the case of most transplant centers nowadays—a negative pre-transplant IFN- γ ELISPT implied an NPV of 91% or higher (ie, a predicted risk of AR of 9% or lower).

Graft Function

Six- and 12-month serum creatinine and eGFR were similar between patients with positive or negative pretransplant IFN- γ ELISPT (weighted mean differences for serum creatinine at 6 and 12 months were as follows: 1.24 mg/dL [−0.72, 3.2], $P = 0.22$ and 0.38 mg/dL [−0.27, 1.04], $P = 0.25$, respectively;

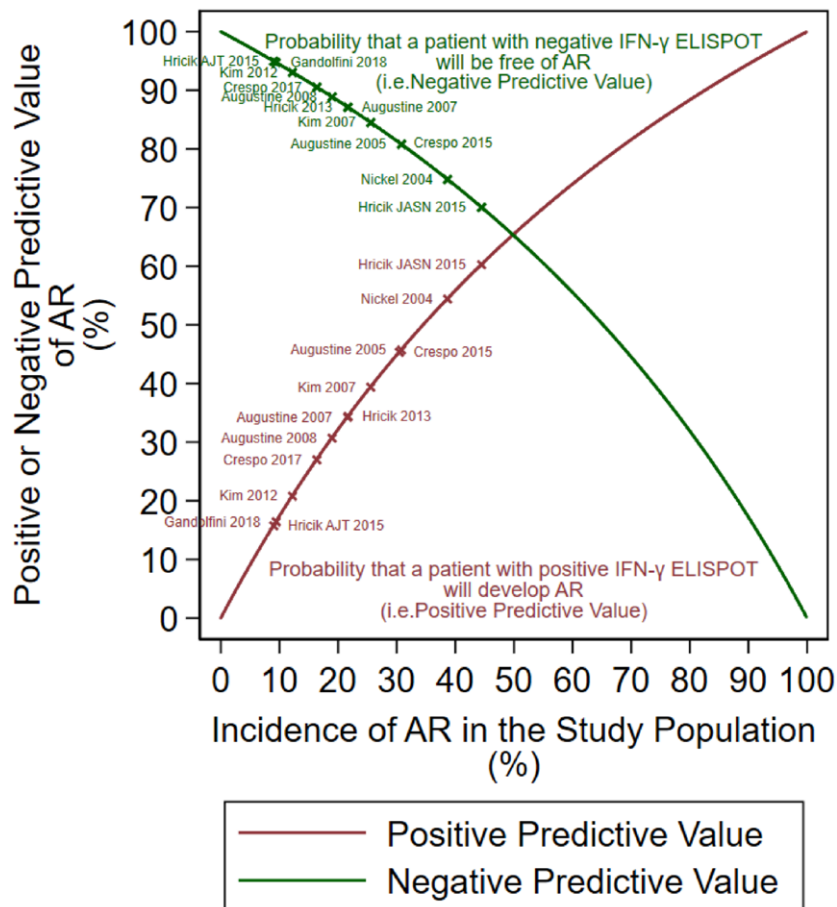


FIGURE 5. Positive and negative predictive values of IFN- γ ELISPOT as a function of the incidence of AR in the study population. Because positive and negative predictive values are not fixed numbers but rather numbers varying with the incidence of AR in the study population, the plot shows on the y-axis how positive or negative predictive values (ie, posttest AR risk) change according to different values of the incidence of AR in the study population (ie, pretest AR risk) reported on x-axis. To visualize how pretest probability changes with the result of IFN- γ ELISPOT, draw a vertical line from a chosen pretest probability of AR on the x-axis until it hits the red curve if IFN- γ ELISPOT is positive, or the green curve if IFN- γ ELISPOT is negative, and read the resulting posttest probability of AR on the y-axis. Cross-marks and associated labels represent the positive and negative predictive values that would result from each study, based on the study-specific incidence of AR of and the overall joined meta-analytic estimate of sensitivity and specificity. AR, acute rejection; IFN- γ ELISPOT, interferon- γ enzyme-linked immunospot.

weighted mean difference for eGFR at 6 and 12 months were as follows: -0.97 [$-2.07, 0.13$], $P = 0.08$ and -0.49 [$-1.27, 0.28$], $P = 0.21$, respectively). However, there was a high heterogeneity between the 2 studies meta-analyzed ($I^2 > 75\%$) (Figures 6 and 7).

TABLE 3.

Positive and negative predictive values of acute rejection

Author	AR incidence	PPV	NPV
Hricik et al ^{11,29}	0.09	0.16	0.95
Gandolfini et al ³¹	0.09	0.16	0.95
Kim et al ¹⁹	0.12	0.21	0.93
Crespo et al ¹⁷	0.16	0.27	0.91
Augustine et al ¹⁶	0.19	0.31	0.89
Augustine et al ¹⁵	0.22	0.34	0.87
Hricik et al ²⁰	0.22	0.34	0.87
Kim et al ²⁰	0.26	0.39	0.85
Crespo et al ¹⁰	0.30	0.45	0.81
Augustine et al ¹⁴	0.31	0.46	0.81
Nickel et al ¹⁸	0.39	0.54	0.75
Hricik et al ^{11,29}	0.44	0.60	0.70

AR, acute rejection; NPV, negative predictive value; PPV, positive predictive value.

Graft and Patient Survival

The very limited number of studies assessing the impact of the pretransplant IFN- γ ELISPOT on graft loss or patient survival did not allow us to perform a meta-analysis on these outcomes. The pooled sensitivity and specificity of pretransplant IFN- γ ELISPOT for both 1-year patient and graft survival prediction are shown in Figures 8 and 9.

DISCUSSION

To the best of our knowledge, this is the first meta-analysis to date investigating the association between pretransplant donor-specific IFN- γ ELISPOT and graft outcomes in kidney transplant recipients. The data show that pretransplant donor-specific IFN- γ ELISPOT is significantly associated with a higher risk of AR. Notably, and although the predictive values of the assay vary depending on the incidence of AR in each study, the NPV of this assay is consistently high (ie, $>90\%$), particularly in low-risk populations (ie, with AR risk $\leq 16\%$). Moreover, and based on our regression analyses on aggregate data, the performance of the assay in predicting graft outcomes seems not to be strongly affected by study population characteristics. We found an apparent trend of specificity to

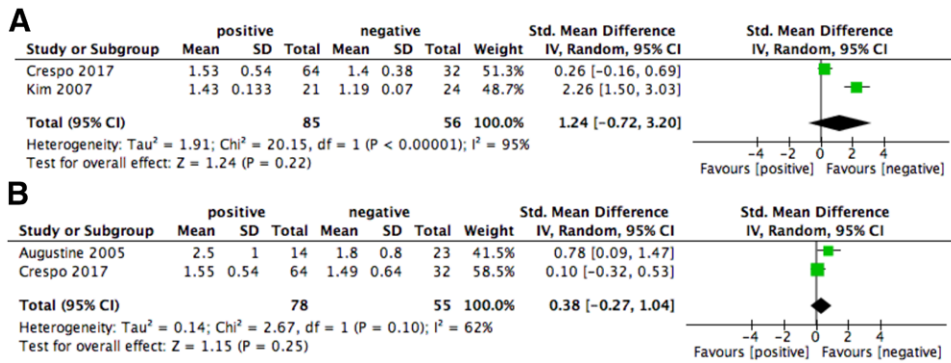


FIGURE 6. Forest plot of serum creatinine (mg/dL) at 6 (A) and 12 months (B). CI, confidence interval; SD, standard deviation.

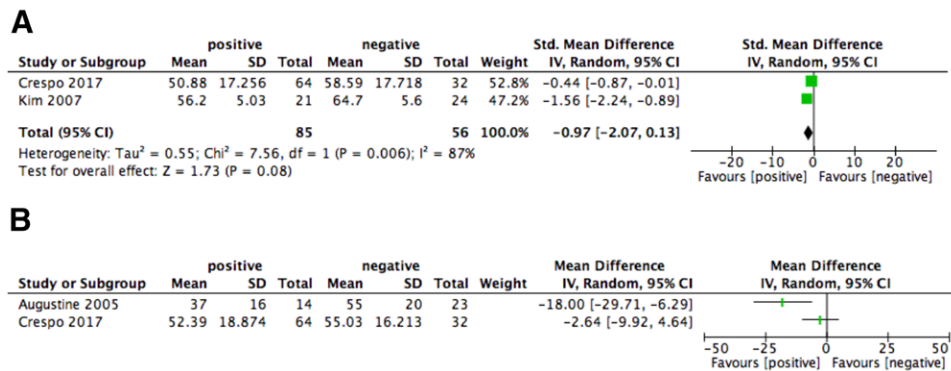


FIGURE 7. Forest plot of eGFR at 6 (A) and 12 months (B). CI, confidence interval; eGFR, estimated glomerular filtration rate; SD, standard deviation.

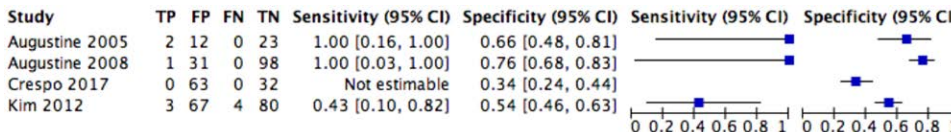


FIGURE 8. Forest plot of graft loss censored for death at 1 year. CI, confidence interval.

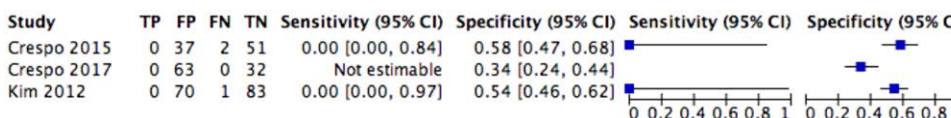


FIGURE 9. Forest plot of analysis of all causes of death at 1 year. CI, confidence interval.

decrease in the years, although we could not infer whether this association was related to unmeasured confounding variables or simply represented a random finding.

Today, the assessment of the immunologic risk in organ transplant recipients before transplantation is exclusively done by the assessment of donor/recipient HLA mismatch and the presence of preformed circulating anti-HLA antibodies,³ with the assumption that the humoral allosensitization also illustrates the allospecific cellular immune response. However, it is well known that cellular memory may occur without humoral activation and is a key factor in initiating and mediating allograft rejection.⁴⁰⁻⁴³ In fact, experimental studies in mice have shown that alloreactive memory T cells play a major pathogenic role in allograft rejection, therefore measuring these cells in a functional manner has been proposed as a strategy to estimate the risk of AR, especially TCMR. The IFN- γ ELISPOT has been developed to measure the frequency

of alloreactive memory IFN- γ -producing T cells at the single cell level.^{44,45} While initial studies showed a clear association between pretransplant IFN- γ ELISPOT positivity and risk of AR,^{8,18} subsequent larger studies challenged this association, especially in subjects receiving T-cell-depleting induction therapy with thymoglobulin.¹⁰⁻¹²

Overall, the present meta-analysis confirms a significant association between pretransplant donor-specific IFN- γ ELISPOT and higher risk of AR, particularly among those patients not receiving T-cell-depleting induction therapy. Though the accuracy to predict AR at the individual level is suboptimal, the high NPV (>90%) of this assay suggests that it could be used to identify patients at low risk of AR. Therefore, pretransplant donor-specific IFN- γ ELISPOT seems to represent an important tool to measure baseline cellular risk stratification in kidney transplant recipients before transplant surgery and eventually allow immunosuppressive

therapy titration accordingly. However, additional biomarkers need to be developed and optimized to better estimate the immunological risk of subjects at the time of kidney transplant. In this regard, as calcineurin inhibitors (CNI) are a key to effectively inhibiting T-cell activation and proliferation, markers accurately identifying patients with suboptimal tacrolimus pharmacokinetics could potentially add value for better patient T-cell immune risk stratification. Likewise, the value of this assay in patients receiving CNI-free immunosuppressive strategies such as costimulation blockade-based regimens needs further investigation.

Our meta-analysis focused on pretransplant donor-specific IFN- γ ELISPOT. Only few studies addressed the utility of serial ELISPOT measurements to detect subclinical AR, predict development of DSA, or graft dysfunction. Data from a prospective cohort study indicate that posttransplant measurements of donor-reactive memory T cells by IFN- γ ELISPOT assay can identify, within kidney transplant patients with stable renal function, patients with minimal risk of having concurrent or future intragraft infiltrates.¹⁷ Likewise, a nonrandomized interventional study showed that prospective evaluation of donor-specific T-cell sensitization by IFN- γ ELISPOT may add important information on the alloimmune state of transplanted patients to be used in daily clinical practice.¹² However, more studies are needed to address whether IFN- γ ELISPOT-driven immunosuppression provides better outcomes than standard, protocol-based clinical management. Two ongoing multicentric, randomized biomarker-driven trials are testing the utility of pretransplant IFN- γ ELISPOT (NCT02540395; NCT03465397) to personalize immunosuppression.

Limitations and Strengths

Our study should be interpreted in light of some limitations. Although we performed an exhaustive search of the literature for pretransplant donor-specific IFN- γ ELISPOT, publication bias cannot be ruled out. It is possible that smaller negative studies have not been published. Another important limitation is the variability in the standard operating procedure used for the IFN- γ ELISPOT. The fact that pretransplant ELISPOT has only been standardized across a relatively small number of centers^{6,7} may, at least in part, explain the heterogeneity across the different studies. Additional limitations include the relatively low number of patients included in the analysis and the use of aggregate as opposed to individual level data; therefore, the statistical power may be inadequate for drawing definitive conclusions about the ability of pretransplant donor-specific IFN- γ ELISPOT to predict in different patient categories such as in patients receiving T-cell depletion induction or among those not receiving CNI-based immunosuppression.

The strengths of this analysis comprise the inclusion of a fairly large number of studies from different continents, which included a wide spectrum of patient categories. Nonetheless, our findings on test performance were reasonably consistent across the different studies.

CONCLUSIONS

In summary, this meta-analysis shows that in low-risk kidney transplant populations a negative pretransplant donor-specific IFN- γ ELISPOT is associated with a very low probability of AR. Although the prediction of AR at the individual level

is suboptimal, this assay could be used, in concert with others, for treatment decision purposes in low-risk populations. Importantly, prospective, randomized, biomarker-driven trials using the IFN- γ ELISPOT assay are warranted.

REFERENCES

- McKenna RM, Takemoto SK, Terasaki PI. Anti-HLA antibodies after solid organ transplantation. *Transplantation*. 2000;69:319–326.
- Lachmann N, Terasaki PI, Budde K, et al. Anti-human leukocyte antigen and donor-specific antibodies detected by luminex posttransplant serve as biomarkers for chronic rejection of renal allografts. *Transplantation*. 2009;87:1505–1513.
- Konvalinka A, Tinckam K. Utility of HLA antibody testing in kidney transplantation. *J Am Soc Nephrol*. 2015;26:1489–1502.
- Enderby C, Keller CA. An overview of immunosuppression in solid organ transplantation. *Am J Manag Care*. 2015;21(1 Suppl):s12–s23.
- Crespo E, Bestard O. Biomarkers to assess donor-reactive T-cell responses in kidney transplant patients. *Clin Biochem*. 2016;49:329–337.
- Ashoor I, Najafian N, Korin Y, et al. Standardization and cross validation of alloreactive ifny ELISPOT assays within the clinical trials in organ transplantation consortium. *Am J Transplant*. 2013;13:1871–1879.
- Bestard O, Crespo E, Stein M, et al. Cross-validation of IFN- γ elispot assay for measuring alloreactive memory/effector T cell responses in renal transplant recipients. *Am J Transplant*. 2013;13:1880–1890.
- Heeger PS, Greenspan NS, Kuhlenschmidt S, et al. Pretransplant frequency of donor-specific, IFN-gamma-producing lymphocytes is a manifestation of immunologic memory and correlates with the risk of posttransplant rejection episodes. *J Immunol*. 1999;163:2267–2275.
- Näther BJ, Nickel P, Bold G, et al. Modified ELISPOT technique—highly significant inverse correlation of post-Tx donor-reactive ifngamma-producing cell frequencies with 6 and 12 months graft function in kidney transplant recipients. *Transpl Immunol*. 2006;16:232–237.
- Crespo E, Lucia M, Cruzado JM, et al. Pre-transplant donor-specific T-cell alloreactivity is strongly associated with early acute cellular rejection in kidney transplant recipients not receiving T-cell depleting induction therapy. *PLoS One*. 2015;10:e0117618.
- Hricik DE, Augustine J, Nickerson P, et al; CTOT-01 consortium. Interferon gamma ELISPOT testing as a risk-stratifying biomarker for kidney transplant injury: results from the CTOT-01 multicenter study. *Am J Transplant*. 2015;15:3166–3173.
- Bestard O, Cruzado JM, Lucia M, et al. Prospective assessment of antidonor cellular alloreactivity is a tool for guidance of immunosuppression in kidney transplantation. *Kidney Int*. 2013;84:1226–1236.
- Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. *Ann Intern Med*. 2013;158:280–286.
- Augustine JJ, Siu DS, Clemente MJ, et al. Pre-transplant IFN-gamma elispots are associated with post-transplant renal function in African American renal transplant recipients. *Am J Transplant*. 2005;5:1971–1975.
- Augustine JJ, Poggio ED, Clemente M, et al. Hemodialysis vintage, black ethnicity, and pretransplantation antidonor cellular immunity in kidney transplant recipients. *J Am Soc Nephrol*. 2007;18:1602–1606.
- Augustine JJ, Poggio ED, Heeger PS, et al. Preferential benefit of antibody induction therapy in kidney recipients with high pretransplant frequencies of donor-reactive interferon-gamma enzyme-linked immunosorbent spots. *Transplantation*. 2008;86:529–534.
- Crespo E, Cravedi P, Martorell J, et al. Posttransplant peripheral blood donor-specific interferon- γ enzyme-linked immune spot assay differentiates risk of subclinical rejection and de novo donor-specific antibodies in kidney transplant recipients. *Kidney Int*. 2017;92:201–213.
- Nickel P, Presber F, Bold G, et al. Enzyme-linked immunosorbent spot assay for donor-reactive interferon-gamma-producing cells identifies T-cell sensitization and correlates with graft function at 6 and 12 months in renal-transplant recipients. *Transplantation*. 2004;78:1640–1646.
- Kim Y, Park KH, Chung BH, et al. Pretransplant IFN- γ ELISPOT assay as a potential screening test to select immunosuppression protocols for patients receiving basiliximab induction therapy. *Transl Res*. 2012;160:230–236.
- Kim SH, Oh EJ, Kim MJ, et al. Pretransplant donor-specific interferon-gamma ELISPOT assay predicts acute rejection episodes in renal transplant recipients. *Transplant Proc*. 2007;39:3057–3060.

21. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group. *JAMA*. 2000;283:2008–2012.
22. Harris RJ, Bradburn MJ, Deeks JJ, et al. meta-analysis: fixed- and random-effects meta-analysis. *Stata J*. 2008;8:3–28.
23. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med*. 2001;20:2865–2884.
24. Harbord RM, Whiting P. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata J*. 2009;9:211–229.
25. Sackett DL, Sackett DL. *Clinical Epidemiology: A Basic Science for Clinical Medicine*. Boston, MA: Little, Brown; 1991.
26. Harbord R, Higgins JPT. Meta-regression in Stata. *Stata J*. 2008;8:493–519.
27. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327:557–560.
28. Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. *Stat Med*. 2003;22:2693–2710.
29. Hricik DE, Formica RN, Nickerson P, et al; Clinical Trials in Organ Transplantation-09 Consortium. Adverse outcomes of tacrolimus withdrawal in immune-quiescent kidney transplant recipients. *J Am Soc Nephrol*. 2015;26:3114–3122.
30. Hricik DE, Poggio ED, Woodside KJ, et al. Effects of Cellular Sensitization and Donor Age on Acute Rejection and Graft Function After Deceased-Donor Kidney Transplantation. *Transplantation*. 2013;95:1254–1258.
31. Gandolfini I, Crespo E, Baweja M, et al. Impact of preformed T-cell alloreactivity by means of donor-specific and panel of reactive T cells (PRT) ELISPOT in kidney transplantation. *PLoS One*. 2018;13:e0200696.
32. Hricik D, Augustine J, Poggio E, et al. Rabbit ATG induction overcomes the negative influence of pretransplant memory T cells on posttransplant GFR: results from the CTOT01 kidney transplant trial [abstract B775]. *Transplantation*. 2014;98:892–893.
33. Crespo E, Lucia M, Melilli E, et al. High pre-transplant frequencies of rATG sensitive anti-donor alloreactive memory/effector T-cells are strongly associated to early acute cellular rejection after kidney transplantation [abstract C1524]. *Transplantation*. 2014;98:323.
34. Bestard O, Cruzado JM, Cassis L, et al. Immune-monitoring of alloreactive memory/effector T-cell alloresponse for selection of CNJ-free immunosuppression in renal transplantation [abstract 1252]. *Am J Transplant*. 2011;11.
35. Hricik D, Heeger P, Padiyar A, et al. Cellular or humoral presensitization increases the risk of delayed graft function and acute rejection in recipients of expanded criteria donor kidneys. In: *Presented at the: American Transplant Congress*, Boston, MA; 2009.
36. Hricik D, Poggio E, Newell K, et al. Early results from the clinical trials in organ transplantation (CTOT-01) trial identify noninvasive markers as correlates of 6-month renal allograft pathology. In: *Presented at the: American Transplant Congress*, Boston, MA; 2009.
37. Oh E, Jung S, Kim Y, et al. Pretransplant donor-specific, and 3rd party-specific interferon-gamma ELISPOT assay predicts acute rejection in renal transplant recipients. In: *Presented at the: American Transplant Congress*, San Diego, CA; 2010.
38. Formica R, Nickerson P, Poggio E, et al. Immune monitoring and tacrolimus (Tac) withdrawal in low risk recipients of kidney transplants—results of CTOT09 [abstract 1436]. *Transplantation*. 2014;98:225.
39. Sackett DL, Haynes RB, Tugwell P, Guyatt G. *Clinical Epidemiology: A Basic Science for Clinical Medicine 2nd rev ed*. Boston, MA: Lippincott Williams and Wilkins; 1991.
40. Amir AL, D'Orsogna LJ, Roelen DL, et al. Allo-HLA reactivity of virus-specific memory T cells is common. *Blood*. 2010;115:3146–3157.
41. Heeger PS. T-cell allorecognition and transplant rejection: a summary and update. *Am J Transplant*. 2003;3:525–533.
42. Issa F, Schiopu A, Wood KJ. Role of T cells in graft rejection and transplantation tolerance. *Expert Rev Clin Immunol*. 2010;6:155–169.
43. Benichou G, Gonzalez B, Marino J, Ayasoufi K, Valujskikh A. Role of memory T cells in allograft rejection and tolerance. *Front Immunol*. 2017;8.
44. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919.
45. Bestard O, Cruzado JM, la Franquesa M, et al. Biomarkers in renal transplantation. *Curr Opin Organ Transplant*. 2010;15:467–473.