

Retrospective Cohort Study

Prediction of delayed graft function using different scoring algorithms: A single-center experience

Magda Michalak, Kristien Wouters, Erik Fransen, Rachel Hellemans, Amaryllis H Van Craenenbroeck, Marie M Couttenye, Bart Bracke, Dirk K Ysebaert, Vera Hartman, Kathleen De Greef, Thiery Chapelle, Geert Roeyen, Gerda Van Beeumen, Marie-Paule Emonds, Daniel Abramowicz, Jean-Louis Bosmans

Magda Michalak, Rachel Hellemans, Amaryllis H Van Craenenbroeck, Marie M Couttenye, Gerda Van Beeumen, Daniel Abramowicz, Jean-Louis Bosmans, Department of Nephrology-Hypertension, Antwerp University Hospital, B-2650 Edegem, Belgium

Kristien Wouters, Department of Medical Statistics, Antwerp University Hospital, B-2650 Edegem, Belgium

Erik Fransen, StatUa Center for Statistics, University of Antwerp, B-2610 Wilrijk, Belgium

Bart Bracke, Dirk K Ysebaert, Vera Hartman, Kathleen De Greef, Thiery Chapelle, Geert Roeyen, Department of Hepatobiliary, Endocrine and Transplantation Surgery, Antwerp University Hospital, B-2650 Edegem, Belgium

Marie-Paule Emonds, Histocompatibility and Immunogenetic Laboratory, Belgian Red Cross-Flanders, 2800 Mechelen, Belgium

Author contributions: Michalak M, Abramowicz D and Bosmans JL designed the study; Michalak M collected and analyzed the data and wrote the first draft of the manuscript; Abramowicz D and Bosmans JL contributed to the analysis of the data and the finalization of the manuscript; Wouters K and Fransen E provided the statistical analysis; Van Beeumen G and Emonds MP contributed to the collection of the patient-related data; Hellemans R, Van Craenenbroeck AH, Couttenye MM, Bracke B, Ysebaert DK, Hartman V, De Greef K, Chapelle T and Roeyen G were responsible for the medical and surgical management of the patients.

Institutional review board statement: This study was approved by the Ethic Committee of the Antwerp University Hospital (Ref. 16/34/339).

Informed consent statement: Since this was a retrospective study, a written informed consent could not be obtained from the patients. However, all patients agreed verbally for the data collection.

Conflict-of-interest statement: This project was an investigator driven study without any support from external organizations. The authors have no conflict of interest.

Data sharing statement: No additional data are available.

Open-Access: This article is an open-access article which was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

Manuscript source: Unsolicited manuscript

Correspondence to: Jean-Louis Bosmans, MD, PhD, Professor of Medicine, Department of Nephrology-Hypertension, Antwerp University Hospital, Wilrijkstraat 10, B-2650 Edegem, Belgium. jeanlouis.bosmans@uantwerpen.be
Telephone: +32-3-8213421
Fax: +32-3-8290100

Received: December 19, 2016

Peer-review started: December 25, 2016

First decision: February 17, 2017

Revised: March 31, 2017

Accepted: May 3, 2017

Article in press: May 5, 2017

Published online: October 24, 2017

Abstract

AIM

To compare the performance of 3 published delayed graft

function (DGF) calculators that compute the theoretical risk of DGF for each patient.

METHODS

This single-center, retrospective study included 247 consecutive kidney transplants from a deceased donor. These kidney transplantations were performed at our institution between January 2003 and December 2012. We compared the occurrence of observed DGF in our cohort with the predicted DGF according to three different published calculators. The accuracy of the calculators was evaluated by means of the c-index (receiver operating characteristic curve).

RESULTS

DGF occurred in 15.3% of the transplants under study. The c index of the Irish calculator provided an area under the curve (AUC) of 0.69 indicating an acceptable level of prediction, in contrast to the poor performance of the Jeldres nomogram (AUC = 0.54) and the Chapal nomogram (AUC = 0.51). With the Irish algorithm the predicted DGF risk and the observed DGF probabilities were close. The mean calculated DGF risk was significantly different between DGF-positive and DGF-negative subjects ($P < 0.0001$). However, at the level of the individual patient the calculated risk of DGF overlapped very widely with ranges from 10% to 51% for recipients with DGF and from 4% to 56% for those without DGF. The sensitivity, specificity and positive predictive value of a calculated DGF risk $\geq 30\%$ with the Irish nomogram were 32%, 91% and 38%.

CONCLUSION

Predictive models for DGF after kidney transplantation are performant in the population in which they were derived, but less so in external validations.

Key words: Delayed graft function; Kidney transplantation; Nomogram; Receiver operating characteristic curve; Risk calculation

© **The Author(s) 2017.** Published by Baishideng Publishing Group Inc. All rights reserved.

Core tip: In this single centre, retrospective study we compared the incidence of observed delayed graft function (DGF) in 247 consecutive kidney transplant recipients with the predicted risk of DGF according to 3 different nomograms. Although the Irish nomogram provided an acceptable predictive value for the global study population, this calculator did not allow to make an accurate prediction of DGF at the individual level. Our study suggests that currently available predictive models for the risk of DGF after kidney transplantation are predictive in the population in which they were derived, but they lose their predictive value in external validations.

Michalak M, Wouters K, Fransen E, Hellemans R, Van Craenenbroeck AH, Couttenye MM, Bracke B, Ysebaert DK, Hartman V, De Greef K, Chapelle T, Roeyen G, Van Beeumen G, Emonds MP, Abramowicz

D, Bosmans JL. Prediction of delayed graft function using different scoring algorithms: A single-center experience. *World J Transplant* 2017; 7(5): 260-268 Available from: URL: <http://www.wjgnet.com/2220-3230/full/v7/i5/260.htm> DOI: <http://dx.doi.org/10.5500/wjt.v7.i5.260>

INTRODUCTION

Delayed graft function (DGF) is classically defined as the need for at least one postoperative dialysis session during the first week after transplantation^[1,2]. This definition has some limitations since the postoperative requirement of dialysis is not standardized and the decision to dialyze is subjective. For this and other reasons, the frequency of DGF varies worldwide between 10% and 40% for deceased donor kidney transplants^[1,3]. DGF leads to prolonged hospitalization, higher cost of transplantation, and increased complexity of management of immunosuppressive drugs^[4-6]. DGF is associated with an increased risk of acute rejection and may negatively impact long-term allograft function and outcome^[7,8].

There are currently neither clinical practice guidelines nor an approved therapy to prevent DGF. In addition, the use of "extended criteria donors" (ECD) and kidneys from donors after cardiac death (DCD), which are associated with a higher incidence of DGF, is rising. The ability to predict DGF at the time of the transplant offer might help in clinical decisions making, such as declining the offer, selecting a recipient who would have a lower DGF risk, or modifying the transplantation strategy. This may include efforts to shorten the cold ischemia time (CIT), or delay the initiation of calcineurin inhibitors (CNIs) under the cover of induction therapy with anti-lymphocyte antibodies, or even to machine-perfuse the kidney.

Recently several DGF-scoring systems have been developed. In 2003, Irish *et al*^[6], using a combination of 16 donor- and recipient-related risk factors known at the time of transplantation, developed a nomogram to predict/quantify the risk of DGF after renal transplantation. In 2010 they refined their previously published model using a more recent data set and adding two risk factors (in total 18) to the analysis (Table 1)^[9]. This predictive model has an area under the receiver operating characteristic curve (ROC AUC) of 0.70, which indicates a good degree of discrimination^[9]. In 2009, Jeldres *et al*^[10] developed a simpler but equally accurate scoring system on 532 patients (6 variables, AUC = 0.74) (Table 1). More recently Chapal *et al*^[11] proposed a predictive score that could be calculated by computing only 5 variables with a ROC AUC of 0.73 (Table 1).

The main aim of our study was to conduct a single-center retrospective analysis of a cohort of 247 adult patients to evaluate the performance of available nomograms to predict DGF in our patients, *i.e.*, in a different population than the one they have been tested in. We also studied separately recipients of standard criteria,

Table 1 Comparison of variables used in different scoring systems

	DGF risk calculator (Irish <i>et al</i> ^[9])	DGFS scoring system (Chapal <i>et al</i> ^[11])	Jeldres scoring system (Jeldres <i>et al</i> ^[10])
Recipient variables			
Recipient BMI	+	+	-
Recipient age	+	-	+
No. of HLA mismatches	+	-	+
Peak PRA (%)	+	-	+
Recipient race	+	-	-
Recipient gender	+	-	-
Duration of dialysis	+	-	-
History of diabetes mellitus	+	-	-
Previous transplantation or blood transfusion	+	-	-
Single or multiple organ transplant	+	-	-
Recipient weight	-	-	+
Donor variables			
Donor age	+	+	+
Duration of CIT	+	+	+
Terminal serum creatinine	+	+	-
Donor weight	+	-	-
Primary cause of death	+	-	-
History of hypertension	+	-	-
Duration of WIT	+	-	-
Type of the donor (living, DCD)	+	-	-
Type of induction therapy	-	+	-

PRA: Panel-reactive antibody; WIT: Warm ischemia time; CIT: Cold ischemia time; DCD: Donation after cardiac death; DGF: Delayed graft function.

extended criteria and donation after cardiac death donors.

MATERIALS AND METHODS

Patient characteristics

From January 1st 2003 to December 31st 2012, 349 renal transplantations were performed at the Antwerp University Hospital. Data were collected from our prospective institutional database and the database of Eurotransplant International Foundation. We excluded 27 pediatric transplants (aged < 18), 16 combined solid organ transplantations in adults (13 pancreases and 3 hearts), 31 transplantations performed with living donors (10.1%), 2 pre-emptive transplantations and 15 machine perfused kidneys. Moreover, we excluded 5 patients because of missing data for CIT. Thus, a total of 253 kidney transplantations from a deceased donor (87% first and 13% re-grafts), performed on 243 patients were considered for study. Six out of those 253 grafts (2%) were lost due to primary non function (PNF). These patients were excluded from further analysis and the final data set comprised 247 transplantations. Recipient and donor characteristics at the time of transplantation are summarized in Tables 2 and 3.

Definition of DGF, PNF and ECD

DGF was defined as the requirement of at least one dialysis within the first 7 d post-transplantation. The duration of DGF was defined as the number of days between the transplantation and the day of the last

dialysis. PNF was defined as the absence of allograft function starting immediately after transplantation, and requiring maintenance dialysis. An ECD was defined as: A donor aged ≥ 60 years, or a donor aged 50-59 years with at least 2 of the following conditions: History of hypertension, terminal serum creatinine level greater than 1.5 mg/dL, or death resulting from a cerebrovascular accident/stroke (CVA).

Post-transplant immunosuppressive therapy

One hundred and sixty-one patients (63.6%) were given an induction with an inhibitor of the IL2-receptor (basiliximab or daclizumab). Ninety-two patients (36.4%) were induced with antithymocyte globulin (ATG). According to our induction immunosuppression protocols ATG was given to immunized patients (peak PRA > 50%), patients of North-African origin, patients with a history of acute rejection during the first year after previous transplantation or in the case of kidney transplantation with ECD or DCD donor kidneys. Most patients (*n* = 244, 96.4%) received a CNI as initial therapy in addition to the treatment with corticosteroids and mycophenolate mofetil. Cyclosporin A was initiated at a starting dose of 2 × 4 mg/kg at post-transplant day 1. Only 7 patients (2.8%) were given mTOR-inhibitors. Two remaining patients (0.8%) [Eurotransplant Senior Program (ESP)] did not receive either medication but only ATG, MMF and prednisolone.

Data collection and DGF risk assessment

Risk factors for DGF included donor^[12-15] and recipient factors known before and at the time of the transplantation

Table 2 Recipient characteristics at the time of transplantation

Age (yr)	50.2 ± 11.9 ²
Origin (%)	
Blacks	4.5
Caucasians	95.5
Gender (%)	
Male	61.9
History of diabetes mellitus (%)	
Yes	16.6
Body mass index (kg/m ²)	25.1 ± 3.8 ²
Pretransplant transfusions (%)	
Yes	38.1
No	56.7
Unknown	5.3
Duration of the pre-transplant renal replacement therapy (mo)	26.7 (16.4-43.5) ¹
Peak panel-reactive antibodies (%)	88.5
≤ 5%	9.5
5%-80%	2
≥ 80%	
Proportion of kidney re-graft (%)	12.6
Total HLA mismatches	3 (2-3) ¹

¹Median with P25-75; ²Mean ± standard deviation (SD).

and were required to calculate the risk of DGF with the DGF risk calculator^[9] (www.transplantcalculator.com/DGF), the Jeldres scoring system (Jeldres *et al.*^[10]) and the DGFS scoring system^[11]. Recipient variables included: Age, gender, race, body mass index (BMI), history of diabetes mellitus, previous transplantation, pretransplant blood transfusion, duration of renal replacement therapy (RRT), the percentage of serum panel-reactive antibodies (peak PRA), and the number of HLA mismatches. Donor variables included: Age, gender, weight, donor type [standard criteria donor (SCD), ECD, DCD], primary cause of death, history of hypertension, duration of cold (CIT) and second warm ischemia time (WIT), and the terminal serum creatinin (mg/dL).

Statistical analysis

The statistical methods were performed and reviewed by Kristien Wouters (Department of Medical Statistics, Antwerp University Hospital, B-2650 Edegem, Belgium) and by Erik Fransen (StatUa Center for Statistics, University of Antwerp, B-2610 Wilrijk, Belgium).

Normality was tested with the Shapiro-Wilk and the Q-Q plot test. Normally distributed data are represented as mean and standard deviation; non-normally distributed data as median with P25 and P75. Categorical data are presented as numbers and percentages. Comparison of predicted DGF probability between DGF positive and negative patients was done by means of the Mann-Whitney *U* test. Receiver operating characteristic (ROC) curves were generated to evaluate the performance of explanatory scoring systems in predicting outcomes. The c-statistic (or AUC = area under ROC curve) was used as a measure of the predictive performance of the studied scoring systems. Additionally, the performance of the 3 nomograms was evaluated using a Hosmer-Lemeshow goodness-of-fit test. All data were analyzed using IBM

Table 3 Donor characteristics at the time of transplantation

Age (yr)	45.1 ± 14.1 ²
Weight (kg)	76.2 ± 16.4 ²
History of hypertension (%)	
Yes	23.1
No	74.5
Unknown	2.4
Terminal serum creatinine (mg/dL)	0.78 (0.61-1.00) ¹
Donor type (%)	
Standard criteria donor	68.8
Extended criteria donor	17
Donation after cardiac death donor	14.2
Primary cause of death (%)	
Cerebrovascular accident/stroke	27.1
Anoxia	8.1
Other	64.8
Cold ischemia time (h)	14 ± 4.7 ²
Second warm ischemia time (min)	32.8 ± 7.9 ²

¹Median with P25-75; ²Mean ± standard deviation (SD).

SPSS statistics (version 21). Statistical significance was predefined as a *P*-value < 0.05. Goodness-of-fit was set at *P* > 0.05 for the Hosmer-Lemeshow test.

RESULTS

DGF occurred in 38 of the 247 transplants under study (15.3%). The mean duration of DGF was 11.3 ± 15.1 d (range 1-71 d). Graft survival at one year was comparable in patients with or without DGF (94.6% vs 93.3% respectively, *P* = ns). However, graft function was significantly inferior in patients with DGF both at 30 d (creatinine clearance according to MDRD formula 31 ± 16 mL/min vs 46 ± 17 mL/min, *P* = 0.001) and at 1 year (42 ± 14 mL/min vs 52 ± 17 mL/min, *P* < 0.001).

Analysis according to the algorithm of Irish *et al.*^[9]

At the population level, the average DGF risk calculated with the DGF risk calculator was 18.5%, which was close to the observed data (DGF rate: 15.3%). The AUC was 0.69 (Figure 1). Figure 2A illustrates the relatively good calibration of the Irish model. The predicted DGF risk and the observed DGF probabilities were close (*P* = 0.74, Hosmer-Lemeshow statistic). The mean calculated DGF risk was significantly different (*P* < 0.0001) between DGF-positive and DGF-negative subjects (Figure 3). However, at the level of the individual patient the calculated risk of DGF overlapped very widely (Figure 3). Indeed, it ranged from 10% to 51% for recipients with DGF and from 4% to 56% for those without DGF. The sensitivity, specificity and positive predictive value of a calculated DGF risk ≥ 30% were 32%, 91% and 38% respectively.

Analysis according to the algorithm of Jeldres *et al.*^[10]

At the population level, the average DGF risk calculated with Jeldres nomogram was 27.9%, which is almost the double of the observed DGF rate (15.3%). The AUC of the ROC curve was poor at 0.54 (Figure 1). The

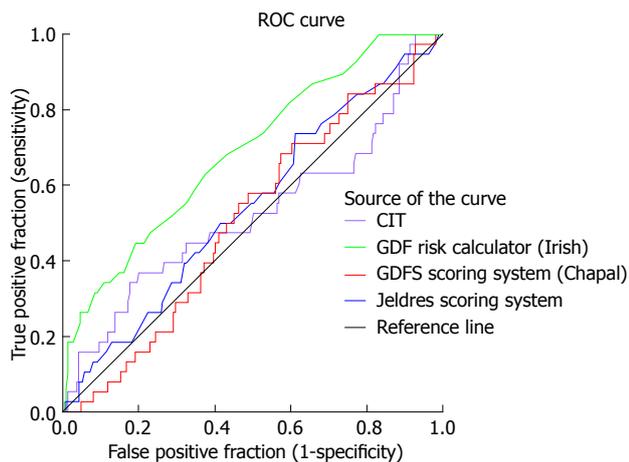


Figure 1 Receiver operating characteristic curves to evaluate the prognostic capacity of cold ischemia time, the delayed graft function risk calculator, the Jeldres scoring system^[10] and the DGFS scoring system^[11] to predict delayed graft function. The cold ischemia time (purple-line): Area under ROC curve (AUC) = 0.52. The DGF risk calculator (green-line) proposed by Irish *et al*^[9]: AUC = 0.69. The scoring system (blue-line) proposed by Jeldres *et al*^[10]: AUC = 0.54. The DGFS scoring system (red-line) proposed by Chapal *et al*^[11]: AUC = 0.51. ROC: Receiver operating characteristic; CIT: Cold ischemia time; DGF: Delayed graft function.

Hosmer-Lemeshow “goodness-of-fit” test demonstrated a significant difference ($P < 0.05$) between the predicted DGF risk and the observed DGF, which indicates that the DGF risk was not well estimated by the Jeldres scoring system (Figure 2B). The calculated risk of DGF showed a wide range of values from 5%-82% in the DGF-group and 3%-83% in the non- DGF-group with a very large overlap between both groups (Figure 4). The sensitivity, specificity and positive predictive value of a calculated DGF risk $\geq 30\%$ was 44.7%, 61.7% and 17.5% respectively.

Analysis according to the algorithm of Chapal *et al*^[11]

The average DGFS value was $-0.48 [(-0.46) \pm 0.76; 95\%CI: (-0.43) - (-0.71)]$ in the DGF positive group and $(-0.48) \pm 0.89; 95\%CI: (-0.46) - (-0.60)$ in the DGF negative group] (Figure 5A), indicating the inability of the Chapal score to predict DGF in our population. The sensitivity, specificity and negative predictive value of a DGFS value $\leq (-0.5)$ were 45.6%, 70.3% and 85.8% respectively. Only 3 patients (1.2%) had a DGFS score ≥ 1.2 , which should in theory point to a high risk of DGF. None of these 3 patients developed DGF (sensitivity and positive predictive value for DGFS score ≥ 1.2 was 0).

The average DGF risk calculated with the DGFS nomogram was 20%. The ROC curve analysis showed a c-index of 0.51 (Figure 1), indicating the absence of any predictive value. There was no difference between the median calculated DGF risk in the DGF-positive and the DGF-negative subjects (Figure 5B). The calibration plot of this model (Figure 2C) showed a significant difference ($P = 0.02$) between the predicted DGF risk and the observed DGF, which indicates that the DGF risk was not well calibrated by the Chapal nomogram. The

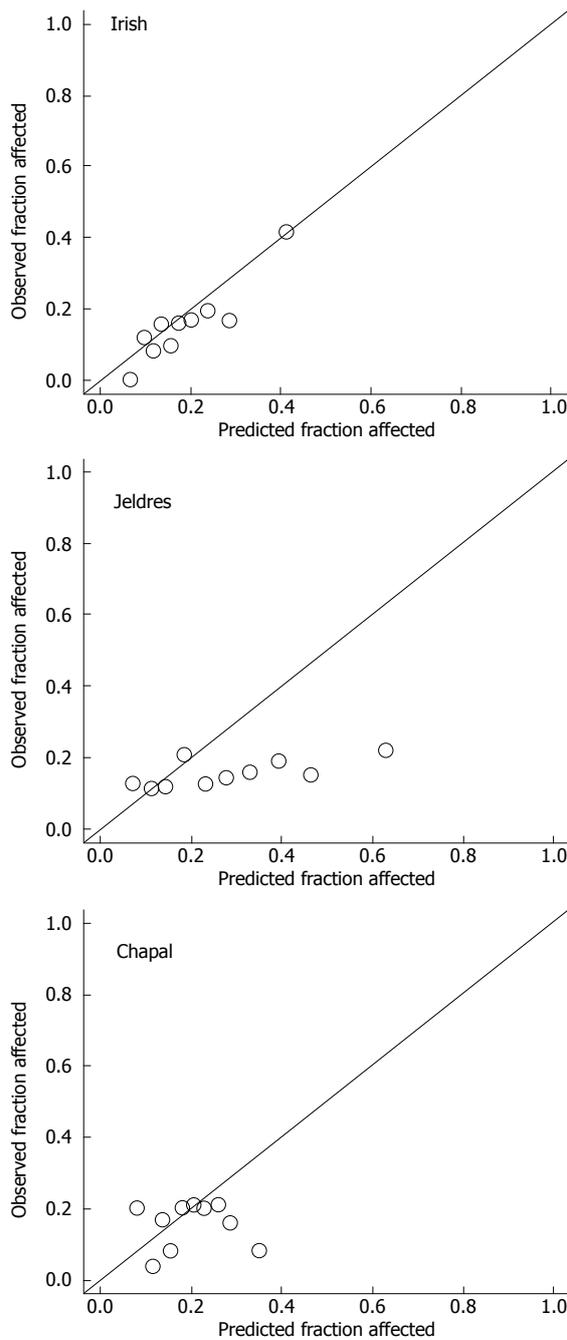


Figure 2 Calibration plot of: The delayed graft function risk calculator (Irish *et al*^[9]), the Jeldres scoring system^[10] and the DGFS scoring system (Chapal *et al*^[11]) to predict delayed graft function. Patients were divided into 10 subgroups (deciles of increased DGF risk), based upon the risk prediction. Each figure plots the mean predicted probability (X-axis) of DGF against the observed prevalence of DGF (Y-axis) (Hosmer-Lemeshow). The P -values were 0.74 for the Irish score, < 0.05 for the Jeldres score and 0.02 for the Chapal score. DGF: Delayed graft function.

sensitivity, specificity and positive predictive value of a calculated DGF risk $\geq 30\%$ were 5.2%, 88% and 8% respectively.

Analysis in the subgroups with a higher risk of DGF

Next, we studied how well the three nomograms can predict DGF in subgroups of patients considered to be at increased risk of DGF such as ECD and DCD donors

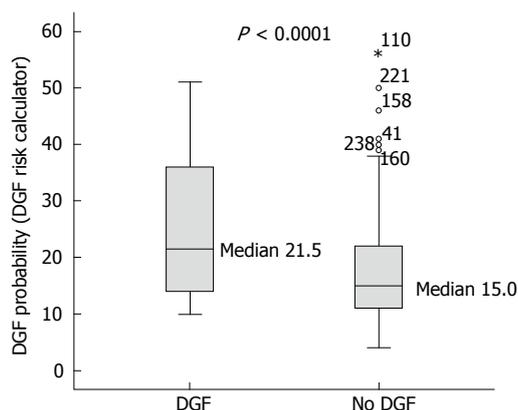


Figure 3 Correlation between the predicted delayed graft function probability according to the delayed graft function risk calculator (Irish *et al*^[9]) and the presence or absence of delayed graft function. DGF: Delayed graft function.

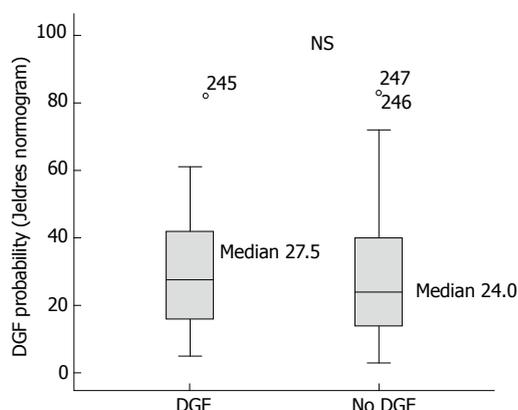


Figure 4 Correlation between the predicted delayed graft function probability according to the Jeldres scoring system (Jeldres *et al*^[10]) and the presence or absence of delayed graft function. DGF: Delayed graft function.

(Table 4). The results presented in Table 4 suggest an acceptable agreement between the observed prevalence of DGF and the Irish DGF score for DCD donors, but not for ECD donors. The DGFS scoring system and the Jeldres scoring system^[10] could not predict DGF in these high-risk groups (Table 4).

DISCUSSION

The first finding from our study is that our mean DGF rate was in the low range (15%), with a stepwise increase according to the risk categories (SCD, ECD, DCD donors). Next, we found that, at a population level, the observed DGF rate and the median calculated DGF risk according to the Irish calculator (16%) were similar. In our study the AUC calculated according to the Irish calculator was 0.69 which is similar to the results obtained in the 2010 Irish model (AUC of 0.70) and indicates an acceptable degree of discrimination. Along this line, the Hosmer-Lemeshow “goodness-of-fit” test demonstrated that the DGF risk was well calibrated by the DGF risk calculator. With regards to the ECD and DCD high-risk groups, there was a good agreement for

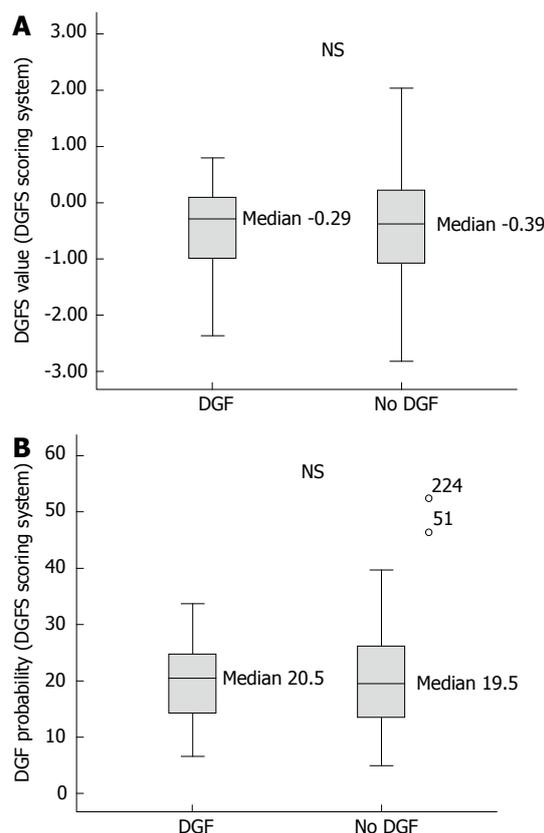


Figure 5 Correlation between the DGFS value (A: Y-axis) and the predicted delayed graft function probability according to the DGFS scoring system (Chapal *et al*^[11]) (B: Y-axis) and the presence or absence of delayed graft function (A and B: X-axis). DGF: Delayed graft function.

DCD but not for ECD. This could be due to the smaller number of patients tested with these conditions in our center. While it appears that the DGF risk calculator can relatively well predict the percentage of DGF in our global study group, it is obvious that we cannot use this tool to take clinical decisions for individual patients. Indeed, as seen in Figure 3, because of the large overlap in DGF risk prediction between patients who developed DGF and those who did not, a high- or low-risk score did not correspond with the presence or absence of DGF. The specificity, sensitivity and positive predictive value of the DGF-risk calculator are too low to help with clinical-decision making regarding the immunosuppressive strategy. This nomogram has been previously tested in Australian^[8], North American^[16] and European^[17] populations, but yielded conflicting results. In the Australian cohort from Kaiser *et al*^[8] the nomogram was applied to 598 deceased donor renal transplantations, and showed a slightly better AUC value of 0.76 with a sensitivity of 74% and a specificity of 71%. Of note, however, no data are given about the overlap between the DGF and no DGF patients in this series, and it is thus difficult to evaluate its predictive value at individual patient level. Moore *et al*^[17] evaluated the nomogram of Irish on 210 United Kingdom patients and showed a similar predictive value with an AUC of 0.71 with a high specificity (95%) but a very poor

Table 4 Observed prevalence vs predicted probability of delayed graft function in the overall population and by risk group

Kidney graft according to donor type	Observed prevalence of DGF (%)	Probability of DGF predicted by the DGF risk calculator (%) (Irish <i>et al</i> ^[9])	Probability of DGF predicted by the DGF scoring system (%) (Chapal <i>et al</i> ^[11])	Probability of DGF predicted by the Jeldres scoring system (%) (Jeldres <i>et al</i> ^[10])
Overall population (n = 247)	15.3	16 ¹ 12-24 ² 0.69 ³	19.7 ¹ 13.6-26 ² 0.51 ³	25 ¹ 14-40 ² 0.54 ³
Standard criteria donor (n = 170)	11.8	14 ¹ 10-20 ² 0.73 ³	20.1 ¹ 14.5-26.4 ² 0.60 ³	21 ¹ 13.7-34.2 ² 0.54 ³
Extended criteria donor (n = 42)	19	19.5 ¹ 14-25 ² 0.39 ³	21.2 ¹ 14.4-27.6 ² 0.34 ³	41.5 ¹ 25.7-60 ² 0.38 ³
Donation after cardiac death (n = 35)	28.6	30 ¹ 18-38 ² 0.65 ³	11.8 ¹ 9.1-20.4 ² 0.58 ³	21 ¹ 8-39 ² 0.64 ³

¹Median; ²P25-P75; ³AUC of the ROC curve. DGF: Delayed graft function.

sensitivity (25%) at a score > 150. They concluded that the utility of the nomogram score in predicting DGF was moderate at best. Grossberg *et al*^[16] showed a poor association between the Irish nomogram and DGF (the average DGF risk in DGF-positive patients was 0.45 ± 0.14 vs 0.40 ± 0.14 in DGF-negatives, P = 0.07) in a US population of 169 patients, but they did not report a c-index.

In 2012, Rodrigo *et al*^[18], used the web-based calculator to predict DGF on 342 European renal transplants. Similar to the Irish group^[9] they found an AUC of 0.71. The reported specificity and sensitivity of a calculated DGF risk ≥ 30% were 75.8% and 51.8% respectively. They concluded, like us, that there was overlap in DGF risk prediction, which limited the utility of the score for individual patients. Finally, a large number of variables are needed to calculate the Irish DGF risk score, which limits its usefulness in daily clinical practice.

For this particular reason, two other independent and easier scoring systems were developed^[10,11]. Jeldres *et al*^[10] developed a more user-friendly nomogram based on the analysis of 6 risk factors. The c-statistic for assessing the predictive ability of Jeldres score for DGF (internal validation) was very similar to the Irish scoring system (AUC of 0.74). However, Chapal *et al*^[11] tried to validate Jeldres score on their patients and showed an inferior predictive capacity of this scoring system to predict DGF (AUC = 0.61). The ROC curve analysis based on our population showed that the predictive utility of the Jeldres scoring system was poor, with a c-index of 0.54. This poor predictive value was confirmed by the Hosmer-Lemeshow “goodness-of-fit” test that showed a bad calibration of this model. The median calculated DGF risk in the DGF-positive group did not differ significantly from the DGF-negative group and there was a large overlap between both groups. Jeldres *et al*^[10] proposed no cut-off to classify patients according to their DGF risk in their original study.

In our study the predictive capacity of the DGFS scoring system from Chapal was poor with an AUC of

0.51. In our population the negative predictive value of the DGFS score was 0.86 which implies that with the DGFS scoring system we can fairly well recognize the patients at a low risk of DGF. In contrast, the threshold for high risk of DGF was clinically useless in our study (none of the patients with DGFS score ≥ 1.2 actually developed DGF). The failure of the DGFS scoring system in the prediction of DGF in our study may be explained by a lower incidence of DGF in our population (15.3% in our study vs 25.5% in the study of Chapal *et al*^[11]). This difference is the consequence of shorter CIT [14 h (range 2.8 to 29.9 h) vs 19.2 h (range 6.0 to 58.6 h)], use of kidneys from younger donors (45.1 years vs 51.9 years) and lower terminal donor serum creatinine (69 µmol/L vs 91 µmol/L) in respectively our study population and in the study by Chapal^[11]. According to these data our center seems to be more stringent in the selection of donors. This could also explain why the algorithm proposed by Chapal *et al*^[11] fails to predict adequately DGF in our population.

There are some limitations to our study. First, the need to dialyse within the first week after the transplantation is an endpoint that could be influenced by several clinical factors (such as for instance heart failure, hyperkalemia...). This can lead to obvious mistakes in the validation of different scoring systems. Second, the sample size in our study is relatively small, particularly when compared to large-population-based transplant registers. Finally, the composition of our study population differs from the initial studies [e.g., 4.5% blacks in our population vs 30.1% blacks in the study of Irish; relatively short CIT in our study (14 ± 4.7 h vs 19.2 ± 7 h in the study of Chapal or 17.8 ± 7.8 h in the study of Irish)]. And finally, according to our induction immunosuppression protocols ATG was de facto given to the patients at increased risk for DGF. The delayed introduction of CNIs could have attenuated the incidence of DGF in our population at risk. Another issue not captured by any scoring system is the policy of peri-operative volemia control, which has been shown to play an important role in the incidence of DGF (Mikhalski

et al^[41]).

In summary, our study suggests that currently available predictive models for the risk of DGF after kidney transplantation are predictive in the population in which they were derived, but they lose their predictive value in external validations. This is not surprising, as none of these scores has been previously rigorously validated in external population of patients. Along this line, there were large variations between centers regarding demographic values (donor age, CIT, proportion of ECD/DCD, *etc.*) explaining why external validation like the one we tried, failed. This means that we still need better predictive tools for the kidney allocation to individual patients, especially those patients who are at high risk of DGF. Currently we are unable to further improve the outcome of a single patient by altering our management on the basis of available scores for the risk of DGF.

COMMENTS

Background

Delayed graft function (DGF) occurs in 10% to 40% of deceased donor kidney transplantations, and leads to prolonged hospitalization, higher costs of transplantation, and increased complexity of management of immunosuppressive drugs. The ability to predict DGF at the time of the transplant offer might help in clinical decision making, such as declining the offer, selecting a recipient who would have a lower DGF risk, or modifying the transplantation strategy. Three predictive scoring systems for DGF were previously developed and published (Irish *et al*, Jeldres *et al* and Chapal *et al*). However, since these scores were not validated in an external study population, we decided to analyse the performance of these three scoring systems in a single centre cohort of 247 consecutive kidney transplant recipients at our institution between 2003 and 2013.

Research frontiers

Three different scoring systems for the prediction of DGF have been developed and validated in the past in respectively well-defined study populations, specific for each study. However, these scoring systems were never validated in an external study population (*i.e.*, different from the initial study population). To explore the validity of these three predictive models, we retrospectively analysed their performance in a cohort of 247 consecutive kidney transplant recipients at our institution.

Innovations and breakthroughs

DGF occurred in 15% of this study population. Only the Irish calculator provided an acceptable level of prediction for DGF with an AUC of the ROC curve of 0.69. However, at the level of the individual patient the calculated risk of DGF overlapped very widely, and therefore this predictive score was not useful in clinical decision making in our study population.

Applications

Based on the reported literature and on our data, we conclude that predictive models for DGF are performant in the population in which they were derived, but these models require additional validation in an external study population.

Terminology

DGF: Delayed graft function; AUC: Area under the curve; ROC: Receiver operator curve; C index: The index of concordance is a "global" index for validating the predictive ability of an algorithm (*e.g.*, for the occurrence of DGF); Nomogram: Is a prediction tool based on information from large numbers of patients. Predictive data are put in a mathematical model that enables to calculate a hypothetical outcome measure.

Peer-review

It is very well-conducted study with some interesting findings, mostly pointed

out that we still cannot predict with accuracy the development of DGF. The study design and method, and statistical analysis were all well-thought and accurately followed throughout the paper.

REFERENCES

- 1 Perico N, Cattaneo D, Sayegh MH, Remuzzi G. Delayed graft function in kidney transplantation. *Lancet* 2004; **364**: 1814-1827 [PMID: 15541456 DOI: 10.1016/S0140-6736(04)17406-0]
- 2 Siedlecki A, Irish W, Brennan DC. Delayed graft function in the kidney transplant. *Am J Transplant* 2011; **11**: 2279-2296 [PMID: 21929642 DOI: 10.1111/j.1600-6143.2011.03754.x]
- 3 Yarlagadda SG, Coca SG, Garg AX, Doshi M, Poggio E, Marcus RJ, Parikh CR. Marked variation in the definition and diagnosis of delayed graft function: a systematic review. *Nephrol Dial Transplant* 2008; **23**: 2995-3003 [PMID: 18408075 DOI: 10.1093/ndt/gfn158]
- 4 Mikhalski D, Wissing KM, Ghisdal L, Broeders N, Touly M, Hoang AD, Loi P, Mboti F, Donckier V, Vereerstraeten P, Abramowicz D. Cold ischemia is a major determinant of acute rejection and renal graft survival in the modern era of immunosuppression. *Transplantation* 2008; **85**: S3-S9 [PMID: 18401260 DOI: 10.1097/TP.0b013e318169c29e]
- 5 Kayler LK, Srinivas TR, Schold JD. Influence of CIT-induced DGF on kidney transplant outcomes. *Am J Transplant* 2011; **11**: 2657-2664 [PMID: 22051325 DOI: 10.1111/j.1600-6143.2011.03817.x]
- 6 Irish WD, McCollum DA, Tesi RJ, Owen AB, Brennan DC, Bailly JE, Schnitzler MA. Nomogram for predicting the likelihood of delayed graft function in adult cadaveric renal transplant recipients. *J Am Soc Nephrol* 2003; **14**: 2967-2974 [PMID: 14569108]
- 7 Pieringer H, Biesenbach G. Risk factors for delayed kidney function and impact of delayed function on patient and graft survival in adult graft recipients. *Clin Transplant* 2005; **19**: 391-398 [PMID: 15877804 DOI: 10.1111/j.1399-0012.2005.00360.x]
- 8 Kaiser MO, Johnson DW. Validation of a nomogram for predicting the likelihood of delayed graft function in Australian adult deceased donor renal transplant recipients. *Nephrology (Carlton)* 2006; **11**: 78 [PMID: 16509938 DOI: 10.1111/j.1440-1797.2006.00532.x]
- 9 Irish WD, Ilsley JN, Schnitzler MA, Feng S, Brennan DC. A risk prediction model for delayed graft function in the current era of deceased donor renal transplantation. *Am J Transplant* 2010; **10**: 2279-2286 [PMID: 20883559 DOI: 10.1111/j.1600-6143.2010.03179.x]
- 10 Jeldres C, Cardinal H, Duclos A, Shariat SF, Suardi N, Capitanio U, Hébert MJ, Karakiewicz PI. Prediction of delayed graft function after renal transplantation. *Can Urol Assoc J* 2009; **3**: 377-382 [PMID: 19829730]
- 11 Chapal M, Le Borgne F, Legendre C, Kreis H, Mourad G, Garrigue V, Morelon E, Buron F, Rostaing L, Kamar N, Kessler M, Ladrrière M, Souillou JP, Launay K, Daguin P, Offredo L, Giral M, Foucher Y. A useful scoring system for the prediction and management of delayed graft function following kidney transplantation from cadaveric donors. *Kidney Int* 2014; **86**: 1130-1139 [PMID: 24897036 DOI: 10.1038/ki.2014.188]
- 12 Port FK, Bragg-Gresham JL, Metzger RA, Dykstra DM, Gillespie BW, Young EW, Delmonico FL, Wynn JJ, Merion RM, Wolfe RA, Held PJ. Donor characteristics associated with reduced graft survival: an approach to expanding the pool of kidney donors. *Transplantation* 2002; **74**: 1281-1286 [PMID: 12451266 DOI: 10.1097/01.TP.0000034060.18738.0B]
- 13 Nyberg SL, Matas AJ, Kremers WK, Thostenson JD, Larson TS, Prieto M, Ishitani MB, Sterioff S, Stegall MD. Improved scoring system to assess adult donors for cadaver renal transplantation. *Am J Transplant* 2003; **3**: 715-721 [PMID: 12780563]
- 14 Schold JD, Kaplan B, Baliga RS, Meier-Kriesche HU. The broad spectrum of quality in deceased donor kidneys. *Am J Transplant* 2005; **5**: 757-765 [PMID: 15760399 DOI: 10.1111/j.1600-6143.2005.00770.x]
- 15 Rao PS, Schaubel DE, Guidinger MK, Andreoni KA, Wolfe RA, Merion RM, Port FK, Sung RS. A comprehensive risk quantification score for deceased donor kidneys: the kidney donor risk index. *Transplantation* 2009; **88**: 231-236 [PMID: 19623019 DOI: 10.1097/TP.0b013e3181ac620b]

- 16 **Grossberg JA**, Reinert SE, Monaco AP, Gohh R, Morrissey PE. Utility of a mathematical nomogram to predict delayed graft function: a single-center experience. *Transplantation* 2006; **81**: 155-159 [PMID: 16436956 DOI: 10.1097/01.tp.0000188621.54448.c8]
- 17 **Moore J**, Tan K, Cockwell P, Krishnan H, McPake D, Ready A, Mellor S, Hamsho A, Ball S, Lipkin G, Borrows R. Predicting early renal allograft function using clinical variables. *Nephrol Dial Transplant* 2007; **22**: 2669-2677 [PMID: 17556406 DOI: 10.1093/ndt/gfm249]
- 18 **Rodrigo E**, Miñambres E, Ruiz JC, Ballesteros A, Piñera C, Quintanar J, Fernández-Fresnedo G, Palomar R, Gómez-Alamillo C, Arias M. Prediction of delayed graft function by means of a novel web-based calculator: a single-center experience. *Am J Transplant* 2012; **12**: 240-244 [PMID: 22026730 DOI: 10.1111/j.1600-6143.2011.03810.x]

P- Reviewer: Hilmi I, Sheashaa HA, Taheri S **S- Editor:** Ji FF
L- Editor: A **E- Editor:** Lu YJ





Published by **Baishideng Publishing Group Inc**
7901 Stoneridge Drive, Suite 501, Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: <http://www.f6publishing.com/helpdesk>
<http://www.wjgnet.com>

