Marked variation in the definition and diagnosis of delayed graft function: a systematic review

Sri G. Yarlagadda, Steven G. Coca, Amit X. Garg, Mona Doshi, Emilio Poggio, Richard J. Marcus and Chirag R. Parikh

Abstract

Background. The term delayed graft function (DGF) is commonly used to describe the need for dialysis after receiving a kidney transplant. DGF increases morbidity after transplantation, prolongs hospitalization and may lead to premature graft failure. Various definitions of DGF are used in the literature without a uniformly accepted technique to identify DGF.

Methods. We performed a systematic review of the literature to identify all of the different definitions and diagnostic techniques to identify DGF.

Results. We identified 18 unique definitions for DGF and 10 diagnostic techniques to identify DGF.

Conclusions. The utilization of heterogeneous clinical criteria to define DGF has certain limitations. It will lead to delayed and sometimes inaccurate diagnosis of DGF. Hence a diagnostic test that identifies DGF reliably and early is necessary. Heterogeneity, in the definitions used for DGF, hinders the evolution of a diagnostic technique to identify DGF, which requires a gold standard definition. We are in need of a new definition that is uniformly accepted across the kidney transplant community. The new definition will be helpful in promoting better communication among transplant professionals and aids in comparing clinical studies of diagnostic techniques to identify DGF and thus may facilitate clinical trials of interventions for the treatment of DGF.

Keywords: delayed graft function; definition; diagnosis

Introduction

Delayed graft function (DGF) is the term used to describe the failure of the transplanted kidney to function immediately after transplantation due to ischemia-reperfusion and immunological injury. It can be considered a form of acute kidney injury post-transplantation and is an important complication of kidney transplantation. The frequency of DGF varies from 4 to 10% in living donor transplants [1] and 5 to 50% in deceased donor kidney transplants [2–5]. DGF complicates post-transplant management, increases morbidity, prolongs patient hospitalization and increases health care costs [6–8]. In addition to the well-known complications of acute kidney injury and dialysis, DGF predisposes the graft to both acute and chronic rejection [9,10], and increases the risk of chronic allograft nephropathy and premature graft loss [11–13]. Despite significant advances in transplantation, the incidence of DGF has not decreased. Though this is partly explained by a higher number of grafts from expanded criteria donors and donation after cardiac death in recent times, there have also been no major therapeutic advances in the field of DGF over the last 20 years. Clinical studies in the area of DGF are difficult to compare as there is no uniformly accepted definition for DGF [1]. In addition, there currently is no effective treatment for DGF. The agents that were successful for the treatment of DGF in the laboratory have been disappointing in the clinical setting [14,15]. This may be due to delayed diagnosis of DGF by current clinical methods that employ dialysis and urine output. The unavailability of standardized definitions of DGF or reliable biomarkers that identify concurrent histopathologic features of DGF contributes to the paucity of therapeutic agents and timely therapeutic intervention. The purpose of this systematic review was to identify all the different definitions of DGF in the literature and current methods available to diagnose DGF.

Methods

This review was conducted and reported in accordance with published guidelines, using a pre-specified protocol [16].
Study eligibility

Our inclusion criteria were (1) original publications of randomized controlled trials, cohort or case-control studies on DGF where the primary aim of the study was to either diagnose DGF, determine the risk factors of DGF or study the effect of DGF on long-term outcomes; (2) studies involving human subjects; (3) studies involving living donor and/or deceased donor transplantation in the adult and/or pediatric patient populations and (4) studies published in English language. We excluded duplicate studies published on the same set of patients.

Finding relevant studies

We screened citations from MEDLINE and EMBASE databases since inception to March 2007. The terms ‘delayed graft function’, ‘renal transplantation’, ‘complications’, ‘biomarkers’ and ‘acute renal failure’ were combined with the terms prognosis, mortality, outcomes and diagnosis. We pilot tested the strategies and modified them to make sure we identified known eligible articles. The eligibility of each citation was evaluated and the full-text article of each citation was retrieved for any citation considered potentially relevant. We complemented the search by searching the Cochrane database of randomized controlled trials, the Science Citation Index on the Web of Science database, reviewing the reference lists from original articles and review articles. The ‘related articles’ feature on Pubmed was also used to identify additional studies. Two reviewers independently (S.G.Y. and S.G.C.) screened the citations and those considered potentially relevant were retrieved for full-text review. They independently evaluated the eligibility of each full-text article, with disagreements resolved by consensus. When duplicate studies involving the same set of patients were encountered, the study with the larger set of patients was included.

Data extraction

Two reviewers (nephrologists with clinical research training), using created forms, independently extracted the following data on all studies meeting eligibility criteria: primary aim of the study, study population (living or deceased donor), total number of patients, sample type (convenience versus consecutive) and whether or not a definition of DGF was mentioned. When the study’s primary aim was to determine the prognosis of DGF, the follow-up time, effect on graft survival, patient survival, acute rejection and kidney function were extracted. When the study’s primary aim was to determine the risk factors of DGF, all the relevant risk factors studied were collected. Duplicate data were reviewed and disagreements were resolved by consensus.

Studies emphasizing diagnosis were evaluated using the STARD quality checklist for studies of diagnostic tests [17]. The STARD criteria consist of a checklist of 25 items for all studies on diagnostic accuracy to ensure that a study report contains a clear description of the inclusion criteria for patients, the testing protocols and the criteria for positivity, as well as an accurate account of subjects included in the study and their results. We utilized all of these items for our validity scoring system and assigned a score to each study we reviewed.

Results

We screened 952 citations and excluded 828 articles based on screening of abstracts. Full-text analysis of the remaining 124 articles resulted in the final inclusion of 88 studies that met our inclusion criteria for this review (Figure 1).

Definitions of DGF

There were 65 studies published between 1984 and March 2007 that mentioned a definition for DGF. There were a total of 18 unique definitions used in the 65 studies. These definitions were commonly related to need for dialysis post-transplant (n = 49, 75%), failure of serum creatinine to decrease (n = 10, 14%) or a combination of these two definitions (n = 6, 11%) (Table 1). While the majority of studies used the need for dialysis after transplant to define DGF, there was marked variation in the time frame used: need for dialysis after transplant (n = 2), need for dialysis in the first week after transplant (n = 45), need for dialysis in the 4 days after transplant (n = 1) [18], need for dialysis in the first 10 days after transplant (n = 1) [19]. One study used a DGF definition that specified that the need for dialysis should be for two or more occasions [3], while the rest of the studies did not specify the number of dialysis sessions needed to be diagnosed as having DGF. The majority of the studies were single-center (n = 53, 81%). The studies were also more likely to be conducted in Europe (n = 37, 57%) as compared to the United States (n = 28, 43%). The majority of the cohorts studied comprised of adults (n = 59, 90%). Two (3%) cohorts were exclusively children [20,21] and five (7%) cohorts comprised both children and adults [3,22,23,5,24].

The number of patients in each study ranged from 12 to 86 682. Only 17 (26%) studies described the study population by race [25–28,4,49,36,37,22,23,3,24]. Forty-nine studies enrolled patients who received kidneys from a deceased donor, 1 study enrolled patients who received kidneys from a living donor [36] and 11 studies included both living and deceased donor transplant recipients [31,5,33,34,21,23,22,21,23,24]. Four studies did not mention the donor type [42–45].

Novel methods for diagnosis of DGF

There were 10 studies where the primary objective was to identify and diagnose DGF using novel objective methods [46,19,47,48,38,45,49,33,35,34]. Methods to diagnose DGF were based on histology of the allograft (n = 3) [38,49,45], urinary and serum biomarkers (n = 4) [48,33,47,50], genetics (n = 1) [34] or the use of imaging (n = 2) [46,19] (Table 2). All the studies were single-center studies except one [33]. The number of patients in the studies ranged from 22 to 98. All but one of the studies were published after the year 2000 [46] that represents the period of increased effort to improve the ability to diagnose DGF. The gold standard used to assess the diagnostic properties of these methods was variable and included the
Delayed graft function in renal transplantation

Fig. 1. Flow chart for selection of studies. (1) Reasons for exclusion are not mutually exclusive; some studies have more than one reason for exclusion. (2) A study may belong to more than one category of DGF (diagnosis, risk factor, prognosis).

following: need for dialysis after transplant [19,33,49,38], biopsy-proven ATN [34,50,46], failure of serum creatinine to decrease by <10% a day for three consecutive days [45] and serum creatinine decreasing by <1.1 mg/dL in the first 5 days after transplant [48] (Table 2). The urine biomarkers IL-18 and NGAL (neutrophil gelatin-associated lipocalin), and the measurement of renal plasma flow had a sensitivity of 85% or higher and an area under the curve of 0.9–0.95 for identifying DGF. Interleukin-6 R levels and staining of manganese superoxide in distal tubules of allograft have a sensitivity of 77% and 76%, respectively. The specificity of these tests ranged from 64 to 88%. Expression of (tumor necrosis factor) TNF-α gene expression in the intra-operative biopsy had an area under the curve of 0.87. All the studies on diagnosis that were included in this systematic review had a score of 7 or greater out of a total score of 25 by the STARD criteria.

Discussion

This systematic review brings to light the ambiguity in the definition for DGF. Eighteen unique definitions to define DGF are used in the medical literature with no uniformly accepted definition. We also describe the performance of 10 novel methods to identify DGF. Currently there is no method to reliably diagnose DGF.

Issues in the development of a uniformly accepted definition of DGF

Measurement issues The majority of the studies included in our review used a dialysis-based definition for DGF. The problem with a dialysis-based definition is that the need for renal replacement therapy (RRT) is subjective and is a clinician-dependent decision [51,52]. Patients may need RRT despite good graft function due to hyperkalemia or volume overload. Thus, the new graft of these patients can be inaccurately classified as having DGF despite having a reasonable glomerular filtration rate (GFR) [1,51]. In contrast, patients with residual urine output are less prone to develop these complications and might not require RRT in the first week. Similarly, patients who are transplanted preemptively prior to initiating RRT may not require dialysis post-transplant despite the fact that their allograft may have a minimal GFR [53].

Consequently, patients with significant residual renal function prior to transplantation can be misclassified as not having DGF despite the allograft having a very low GFR. Given the fact that dialysis requirements may not closely correlate with allograft dysfunction, a dialysis-based definition for DGF appears to be inadequate.

Using urine output-based definitions poses a difficulty in patients with significant urine output before transplant. In these patients, it is not possible to differentiate urine output from the native and transplanted kidney. Patients with significant urine output will be classified as not having DGF if the urine output alone is used as a criterion. Similarly, some patients with severe tubular injury might have a preserved urine output (non-oliguric renal failure) in the presence of inadequate renal clearance. For example, even when the GFR falls to 7 L/day (equal to 5 mL/min), the urine output may still be relatively normal at 1–2 L/day.

Creatinine-based definitions are prone to misclassification and require at least 48 h to elapse before a diagnosis of DGF can be made. As a large amount of renal mass can be lost without appreciable changes in serum creatinine and patients are not in a steady state, post-transplant interpretation of allograft function based on serum creatinine can be difficult [54,55]. Also, the frequently performed hemodialysis session just prior to surgery in order to optimize the medical status of the patient may falsely lower the post-transplant creatinine and may masquerade the presence of DGF. Consequently, serum creatinine alone is not ideal to classify DGF.
Table 1. Different definitions used for DGF

<table>
<thead>
<tr>
<th>Definitions</th>
<th>Number of studies</th>
<th>N (no. of patients)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dialysis-based definitions</td>
<td>41</td>
<td>259 251</td>
<td>[60–62,34,63,64,26,65,66,27,87,67,32,68.2,33,22,69,70,43,40,71,14,13,25,21,42,12,72,23,73,74,5,75,20,76,24,3,77–80]</td>
</tr>
<tr>
<td>Need for dialysis in the first week after transplant</td>
<td>2</td>
<td>760</td>
<td>[56,38]</td>
</tr>
<tr>
<td>Need for dialysis in the first week after transplant once</td>
<td>2</td>
<td>737</td>
<td>[79,29]</td>
</tr>
<tr>
<td>Absence of life-sustaining renal function that requires dialysis on</td>
<td>1</td>
<td>41</td>
<td>[19]</td>
</tr>
<tr>
<td>Absence of life-sustaining renal function that requires dialysis in</td>
<td>1</td>
<td>547</td>
<td>[3]</td>
</tr>
<tr>
<td>Absence of life-sustaining renal function that requires dialysis in</td>
<td>1</td>
<td>319</td>
<td>[81]</td>
</tr>
<tr>
<td>Absence of life-sustaining renal function that requires dialysis in</td>
<td>1</td>
<td>263</td>
<td>[18]</td>
</tr>
<tr>
<td>Creatinine-based definitions</td>
<td>5</td>
<td>1471</td>
<td>[9,45,39,82,35]</td>
</tr>
<tr>
<td>Serum creatinine increased or remained unchanged or decreased &lt;10%/day</td>
<td>5</td>
<td>1471</td>
<td>[9,45,39,82,35]</td>
</tr>
<tr>
<td>Day 2 &lt; 1000 mg</td>
<td>2</td>
<td>401</td>
<td>[31,83]</td>
</tr>
<tr>
<td>Serum creatinine &gt; 2.5 mg/dL on Day 7 or the need for post-transplant</td>
<td>1</td>
<td>99</td>
<td>[44]</td>
</tr>
<tr>
<td>Time required for the kidney to reach Crcl &gt; 10 mL/min greater than 1 week.</td>
<td>1</td>
<td>843</td>
<td>[11]</td>
</tr>
<tr>
<td>Failure of creatinine to decline in the first 48 h in the absence of</td>
<td>1</td>
<td>291</td>
<td>[41]</td>
</tr>
<tr>
<td>Delays in diagnosis</td>
<td>More importantly, the shortcoming of all of the definitions used thus far for DGF allow one to make a diagnosis of DGF only after many hours or days have elapsed. This is a disadvantage because animal studies suggest that treatment of acute kidney injury within hours of ischemia-reperfusion injury may be necessary. In animal models of DGF, several therapies applied within hours of induced ischemia-reperfusion injury successfully attenuated renal function after transplant should be referred to as DGF only after other entities that present as early graft dysfunction are ruled out or unlikely. The current definitions do not allow us to distinguish DGF from other causes of graft dysfunction. Newer definitions or diagnostic tests should incorporate these subtleties so that patients are correctly classified as DGF (Figure 2).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The most promising newer methods to diagnose DGF were several methods to rapidly diagnose DGF in the literature. We identified a method that could make the diagnosis immediately after transplantation and should not rely on the need for dialysis or urine output to make the diagnosis. The ideal method should be able to diagnose DGF immediately after transplantation and should not rely on the need for dialysis or urine output to make the diagnosis. We identified several methods to rapidly diagnose DGF in the literature. The most promising newer methods to diagnose DGF were urine biomarkers (IL-18 and NGAL) and the measurement of systemic and renal vein eicosanoids [47]. The advantages of these biomarkers are that it is noninvasive to obtain the urine or blood and also these biomarkers are elevated at early time points after renal injury (within a few hours). However, these biomarkers are still in their initial stages of testing and not yet ready for clinical practice.

**Newer techniques for diagnosing DGF**

A more rapid method to identify DGF clinically is needed. The ideal method should be able to diagnose DGF immediately after transplantation and should not rely on the need for dialysis or urine output to make the diagnosis. We identified several methods to rapidly diagnose DGF in the literature. The most promising newer methods to diagnose DGF were

<table>
<thead>
<tr>
<th>First author</th>
<th>Year of publication</th>
<th>Description of the test</th>
<th>Test characteristics</th>
<th>N</th>
<th>Gold standard/definition used</th>
<th>STARD score (max. score 25)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preidler [46]</td>
<td>1996</td>
<td>Gadolinium-enhanced MR imaging</td>
<td>Not reported</td>
<td>24</td>
<td>Allograft biopsy showing ATN and no rejection</td>
<td>17</td>
</tr>
<tr>
<td>Freedland [19]</td>
<td>2001</td>
<td>EPRF $&lt; 210$ mL/min/1.7 m²</td>
<td>Sensitivity 100%, specificity 69%</td>
<td>41</td>
<td>Need for dialysis within the first 10 days</td>
<td>11</td>
</tr>
<tr>
<td>Norio [47]</td>
<td>2001</td>
<td>Systemic and renal vein eicosanoids</td>
<td>Not reported</td>
<td>62</td>
<td>Failure of serum creatinine to decrease and the need for dialysis during the first week</td>
<td>7</td>
</tr>
<tr>
<td>Parikh [48]</td>
<td>2004</td>
<td>Urine IL-18 in first 24 h &gt;500 pg/mg</td>
<td>Sensitivity 85%, specificity 88% AUC 0.95</td>
<td>22</td>
<td>Serum creatinine decreasing by $&lt; 1$ mg/dL/day for three consecutive days in the first 5 days after transplant or the need for hemodialysis in the first week post-transplantation and a urine sediment characteristic of ATN, i.e. muddy brown broad granular casts and renal tubular epithelial cells</td>
<td>13</td>
</tr>
<tr>
<td>Avihingsanon [38]</td>
<td>2005</td>
<td>TNF-α gene expression in intra-operative biopsy</td>
<td>$R^2 = 0.68 (P &lt; 0.001) AUC 0.87$</td>
<td>75</td>
<td>Need for dialysis during the first week post-transplantation in the absence of AR, vascular complications, or urinary tract obstruction</td>
<td>8</td>
</tr>
<tr>
<td>Boom [45]</td>
<td>2005</td>
<td>Presence of manganese superoxide in distal tubules of renal allograft</td>
<td>Sensitivity 76% (for the absence of DGF)</td>
<td>40</td>
<td>Failure of the graft to reduce serum creatinine concentration by $&gt; 10$% over three consecutive days for more than 1 week after transplantation</td>
<td>11</td>
</tr>
<tr>
<td>Mishra [49]</td>
<td>2006</td>
<td>NGAL staining in an allograft biopsy</td>
<td>Correlation between NGAL staining and post-operative creatinine $R = 0.86$</td>
<td>25</td>
<td>Need for dialysis in the first week after transplant</td>
<td>12</td>
</tr>
<tr>
<td>Parikh [33]</td>
<td>2006</td>
<td>Urine NGAL on Day 0 &gt; 1000 ng/mg and IL-18 on Day 0 &gt; 500 pg/mg</td>
<td>Sensitivity 90%, specificity 83%, AUC 0.9</td>
<td>53</td>
<td>Need for dialysis in the first week after transplant.</td>
<td>15</td>
</tr>
<tr>
<td>Sadeghi [50]</td>
<td>2006</td>
<td>Pre-transplant soluble IL-6R of 35 000 pg/mL</td>
<td>Sensitivity 77%, Specificity 64%*</td>
<td>98</td>
<td>Biopsy-proven ATN</td>
<td>9</td>
</tr>
<tr>
<td>Mehta [34]</td>
<td>2006</td>
<td>CALCa hypermethylation</td>
<td>Trend toward increased aberrant hypermethylation in patients with ATN versus rejection</td>
<td>25</td>
<td>Biopsy-proven ATN</td>
<td>13</td>
</tr>
</tbody>
</table>

TNF: tumor necrosis factor; AUC: area under the curve; AR: acute rejection; NGAL: neutrophil gelatin-associated lipocalin; IL: interleukin; ATN: acute tubular necrosis; MR: magnetic resonance; EPRF: estimated renal plasma flow; CALCa: calcitonin gene.

*A UAC mentioned as significant but value not given.

failure [14]. In contrast, similar therapies in humans, applied several days after the injury (due to the delayed diagnosis of DGF), led to disappointing results [57,15]. As these therapies could have serious adverse effects, it is important that the diagnosis of DGF be made with certainty, thereby sparing the exposure to patients that do not have DGF. Hence, a compelling need to make an early diagnosis of DGF within a few hours after surgery exists. This will allow administration of promising therapies at a time when the injured kidney may still derive benefits.
can no longer be safely administered to patients who have undergone a kidney transplant for end-stage renal disease [58,59]. The presence of manganese superoxide [45], TNF-α gene expression [38] or NGAL staining intensity on the renal allograft [49] requires a kidney biopsy to be performed, which has its associated risks. It has no advantage over a kidney biopsy alone that is the current diagnostic test of choice to confirm DGF.

**Improving the diagnosis of DGF**

A limitation of all of these newer diagnostic tests is that the accuracy of a diagnostic test is assessed by comparing them to a gold standard. Unfortunately, there is no such gold standard in the diagnosis of DGF. The different definitions of DGF make comparing the sensitivity and specificity of one test against another difficult. Hence, we are in need of a unifying definition that would serve as a gold standard and can be universally used across all the research that involves identification of DGF. This would allow evolution of a test that would identify DGF.

Although a specific DGF definition is not well supported in the literature, the authors propose the following schema. First, DGF is present only after entities of early graft dysfunction that are not related to ischemia-reperfusion injury are ruled out or unlikely. Specifically, it is the opinion of the authors that DGF could be reasonably defined as a combination of dialysis need (preferably more than once) or creatinine reduction ratio of less than perhaps 25% (or lower) within the first 48 h post-transplant. Adding creatinine reduction ratio to the definition allows an objective and quantitative diagnosis after transplantation. We think that defining DGF this way, by creatinine reduction ratio, will enable early detection of patients who are able to avoid dialysis in the first few days despite poor graft function. This definition could serve as the gold standard for the definition of DGF and aid in the development of a reliable and objective diagnostic test for DGF.

Ultimately, it would be ideal to diagnose DGF with a technique such as urine or serum biomarkers. This test should be able to identify DGF early and be highly sensitive so that even mild cases of DGF are identified and specific so that patients with other causes of graft dysfunction are not wrongly labeled as DGF. Lastly, such a method should be validated with long-term outcomes of DGF. This situation is akin to the role of troponin in the management of acute myocardial infarction (AMI). Waiting for dialysis-requiring DGF to institute therapy is analogous to waiting until left-ventricular dysfunction and congestive heart failure have occurred after AMI. However, because of the implementation of serum troponin into the diagnostic schema for AMI, the management of AMI has been revolutionized by rapid institution of therapies after the identification of elevated serum troponin concentrations, and the morbidity and mortality from AMI have declined over the past 20 years. It is with this grand hope that we believe an effective therapy for DGF can be discovered and operationalized in the near future once we improve our ability to diagnose DGF early and with precision.

**Limitations**

Our study has some limitations that should be considered. Although we performed an exhaustive search of the literature for DGF studies, it is possible that smaller studies of
post-transplant outcomes that mention DGF in some capacity but not as a primary aim were published but missed by our search. Furthermore, none of the studies of the newer techniques to diagnose DGF have been replicated; thus the reliability of these results is currently uncertain.

Conclusion

In conclusion, this systematic review highlights the heterogeneity of the available and published definitions for DGF and the paucity of reliable ways to identify DGF. Disappointing results with agents for the treatment of DGF that were promising in the laboratory may be due to multiple factors, but the inability of the current methods available to diagnose DGF early and accurately likely contribute. These agents are probably administered when it is too late and little benefit could be achieved. Misclassification with the current definitions of DGF may result in the treatment of many patients who actually do not have DGF and lack of treatment for patients who do have DGF. Hence, we are in need of a unifying definition of DGF that likely involves serum and urine biomarkers in order to improve the field of transplant nephrology.

Acknowledgements. Drs Yarlagadda and Coca are supported by NIH training grant T32DK07276. Dr Parikh was supported by the career development award from NIDDK (K23-DK064689) and ‘The Patrick and Catherine Weldon Donoghue Medical Research Foundation Award’.

Conflict of interest statement. We have no conflict of interest. Part of this work was presented as oral communication at the American Transplant Congress, San Francisco, May 2007.

References


37. Kruger B, Schnitzbauer AA, Boger CA et al. Pretransplant calcium levels have no predictive value for delayed graft function, long-term graft function, cardiovascular events, or graft and patient survival in renal transplantation. *Transplant Proc* 2006; 38: 697–700


Delayed graft function in renal transplantation

75. Howard RJ, Pfaff WW, Brunson ME et al. Delayed graft function is associated with an increased incidence of occult rejection and results in poorer graft survival. Transplant Proc 1993; 25: 884


Received for publication: 29.12.07
Accepted in revised form: 28.2.08