January 22, 2013

Dear Centers for Medicare & Medicaid Services,

I am writing to make a formal request for a new national coverage determination (NCD) for the AlloMap gene expression test that is provided by XDx Expression Diagnostics. I am formally requesting that Medicare choose not to pay a reimbursement for this test, which is an inferior diagnostic test. I will make this case using the publically available evidence below. The AlloMap gene test falls under the benefit category: diagnostic test. The AlloMap gene expression test is designed to identify individuals in need of a heart biopsy to identify acute cellular rejection in a heart transplant population. As such, this test is useful in only a small percentage of Medicare patients. It was approved by the FDA (510(k) Number: k073482) in 2008.

From the FDA file k073482, the indications for use are as follows:

> Indication(s) for use:
> AlloMap Molecular Expression Testing is an In Vitro Diagnostic Multivariate Index assay (IVDMIA) test service, performed in a single laboratory, assessing the gene expression profile of RNA isolated from peripheral blood mononuclear cells (PBMC). AlloMap Testing is intended to aid in the identification of heart transplant recipients with stable allograft function who have a low probability of moderate/severe acute cellular rejection (ACR) at the time of testing in conjunction with standard clinical assessment.
>
> Indicated for use in heart transplant recipients:
> - 15 years of age or older
> - At least 2 months (≥55 days) post-transplant

Traditionally, the surveillance of heart transplant recipients for acute cellular rejection (ACR) and antibody mediated rejection (humoral, AMR) has been performed through the interactions of cardiologists and pathologists. Cardiologists take 3-5 small pieces of heart tissue from the right ventricular septum of the heart by the use of a bioptome in a predefined surveillance routine. These pieces are processed, sectioned and evaluated by a pathologist for the evaluation of ACR and AMR. A grading criteria, made under the auspices of the International Society for Heart and Lung Transplantation, is used internationally to grade biopsies for ACR (Stewart S et al, JHLT, 2005). In the current grading scheme, heart biopsies are graded from 0 to 3R indicating no cellular rejection, mild rejection, moderate rejection and severe rejection. Generally, clinicians are instructed to treat grades 2R and 3R with higher immunosuppression and to maintain their patients on their current drug regimens for any grade 0R or 1R rejection. The heart biopsy remains the standard to make the diagnosis of ACR.

The heart biopsy is also used to make the pathologic diagnosis of AMR, generally by observing positive complement staining for the complement split product C4d. This finding is used by the cardiologist along

with the clinical picture to initiate specific therapies aimed at reducing antibody levels including the use of plasmapheresis. There is no other method that can be used to make the pathologic diagnosis of AMR.

Despite being the standard of care for post-transplant cardiac surveillance, the endomyocardial biopsy is invasive, not without a small risk of life-threatening complications, potentially painful, expensive, and labor intensive. The possibility of a non-invasive test, such as a blood test, to supplant the endomyocardial biopsy would be seen as a significant improvement over the current state of the art.

The origins of the AlloMap test began in 2004. Horwitz et al described a whole-blood expression profile experiment in which they identified 91 transcripts that were differentially expressed between 7 individuals with biopsy proven acute cellular rejection and 7 individuals without acute cellular rejection. By cluster analysis, they could discern an expression profile that predicted acute cellular rejection and one that predicted no rejection (Horwitz et al, Circulation, 2004).

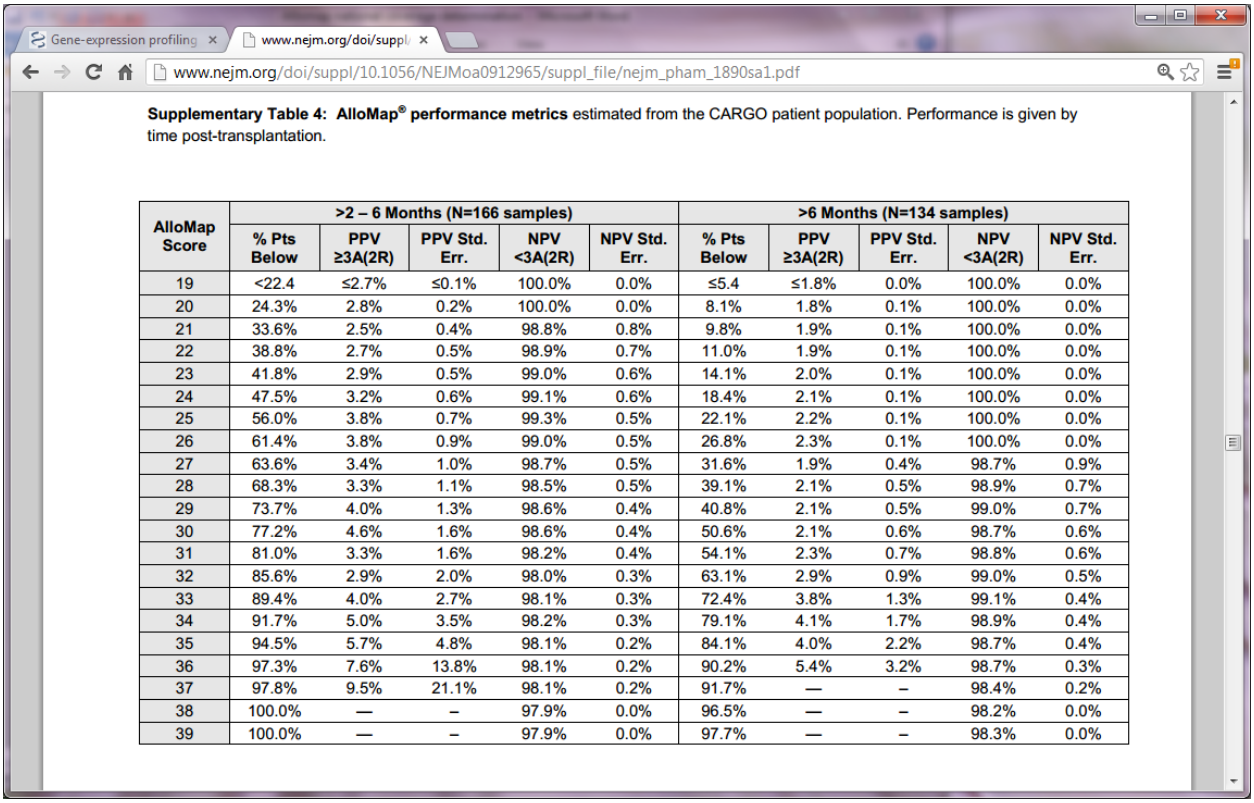
This initial expression experiment was followed by the CARGO trial that was supported by XDx Inc (Deng MC, AJT, 2006). The CARGO trial was a multicenter trial with both a gene discovery and validation arm. The gene discovery was performed using peripheral blood mononuclear cells from 98 patients (285 samples) on a 7,370 gene array. Validation was performed using PCR training. The study identified a 20 gene panel (11 informative and 9 control genes) that could be graded on a 0-40 scale depending on the expression of each gene. The panel was used to distinguish between grade 0 and grades 2R and 3R where the authors reported a positive predictive value (PPV) of 6.8% and a negative predictive value (NPV) of 99.6% in individuals ≥1 year after transplantation with an AlloMap score of 30 or less (out of the 40 point scale). The authors suggested that this test be performed on all appropriate patients and only patients with a score >30 should undergo a followup biopsy. It is important to note that the goal of the project was to find a blood based test to predict patients with acute cellular rejection. However, this 20 gene panel had no PPV and only a strong NPV between the ends of the ACR spectrum. Thus it does not supplant the endomyocardial biopsy as a means to identify acute cellular rejection. It was also a poorer predictor within the first year after transplantation. The panel was not found to have any predictive capabilities for AMR. The authors had made the case that the strong negative predictive value of the test could be an indication to reduce the number of endomyocardial biopsies required post-transplant, which is a reasonable interpretation of the data.

In 2007, Pham MX et al reported on the proposed IMAGE trial, which was designed as a follow up to the CARGO trial. The study was designed as a noninferiority trial of 505 adults who were all at least 1 year post heart transplantation. The primary end points of the study were to be “a composite of: (a) ≥25% proportional decrease in LVEF on echocardiography relative to the enrollment value; (b) any rejection with hemodynamic compromise; and (c) all-cause mortality. The secondary and exploratory end-points include: number of biopsies performed; number of biopsy related complications; health-care resource utilization; and QOL associated with each rejection surveillance strategy.” The goal was to prove that the AlloMap test was not inferior in any of these primary end points and could be met with greater patient satisfaction (having blood drawn versus an invasive procedure).

In 2010, the IMAGE group reported their findings in the New England Journal of Medicine (Pham MX et al NEJM 2010). The group reported a successful noninferiority trial in which there were no statistically significant differences in composite primary outcome or 2-year death rates from any cause. They also reported fewer biopsies performed in the AlloMap group and increased patient satisfaction in the AlloMap group.

However, these sterling findings have almost nothing to do with the quality of the AlloMap test. The high negative predictive value is essentially based upon the reality that most patients (>96%) who undergo a routine endomyocardial biopsy after the 1st year of transplantation will not be found to have treatable rejection (grades 2R and 3R).

If one focuses specifically on the analysis of the AlloMap test performance, one can see that it has no intrinsic predictive capability. The most obvious example of this is from data found in the supplementary figure 4 of the Pham MX NEJM paper.



Supplementary Table 4: AlloMap® performance metrics estimated from the CARGO patient population. Performance is given by time post-transplantation.

AlloMap Score	>2 – 6 Months (N=166 samples)					>6 Months (N=134 samples)				
	% Pts Below	PPV ≥3A(2R)	PPV Std. Err.	NPV <3A(2R)	NPV Std. Err.	% Pts Below	PPV ≥3A(2R)	PPV Std. Err.	NPV <3A(2R)	NPV Std. Err.
19	<22.4	≤2.7%	≤0.1%	100.0%	0.0%	≤5.4	≤1.8%	0.0%	100.0%	0.0%
20	24.3%	2.8%	0.2%	100.0%	0.0%	8.1%	1.8%	0.1%	100.0%	0.0%
21	33.6%	2.5%	0.4%	98.8%	0.8%	9.8%	1.9%	0.1%	100.0%	0.0%
22	38.8%	2.7%	0.5%	98.9%	0.7%	11.0%	1.9%	0.1%	100.0%	0.0%
23	41.8%	2.9%	0.5%	99.0%	0.6%	14.1%	2.0%	0.1%	100.0%	0.0%
24	47.5%	3.2%	0.6%	99.1%	0.6%	18.4%	2.1%	0.1%	100.0%	0.0%
25	56.0%	3.8%	0.7%	99.3%	0.5%	22.1%	2.2%	0.1%	100.0%	0.0%
26	61.4%	3.8%	0.9%	99.0%	0.5%	26.8%	2.3%	0.1%	100.0%	0.0%
27	63.6%	3.4%	1.0%	98.7%	0.5%	31.6%	1.9%	0.4%	98.7%	0.9%
28	68.3%	3.3%	1.1%	98.5%	0.5%	39.1%	2.1%	0.5%	98.9%	0.7%
29	73.7%	4.0%	1.3%	98.6%	0.4%	40.8%	2.1%	0.5%	99.0%	0.7%
30	77.2%	4.6%	1.6%	98.6%	0.4%	50.6%	2.1%	0.6%	98.7%	0.6%
31	81.0%	3.3%	1.6%	98.2%	0.4%	54.1%	2.3%	0.7%	98.8%	0.6%
32	85.6%	2.9%	2.0%	98.0%	0.3%	63.1%	2.9%	0.9%	99.0%	0.5%
33	89.4%	4.0%	2.7%	98.1%	0.3%	72.4%	3.8%	1.3%	99.1%	0.4%
34	91.7%	5.0%	3.5%	98.2%	0.3%	79.1%	4.1%	1.7%	98.9%	0.4%
35	94.5%	5.7%	4.8%	98.1%	0.2%	84.1%	4.0%	2.2%	98.7%	0.4%
36	97.3%	7.6%	13.8%	98.1%	0.2%	90.2%	5.4%	3.2%	98.7%	0.3%
37	97.8%	9.5%	21.1%	98.1%	0.2%	91.7%	—	–	98.4%	0.2%
38	100.0%	—	–	97.9%	0.0%	96.5%	—	–	98.2%	0.0%
39	100.0%	—	–	97.9%	0.0%	97.7%	—	–	98.3%	0.0%

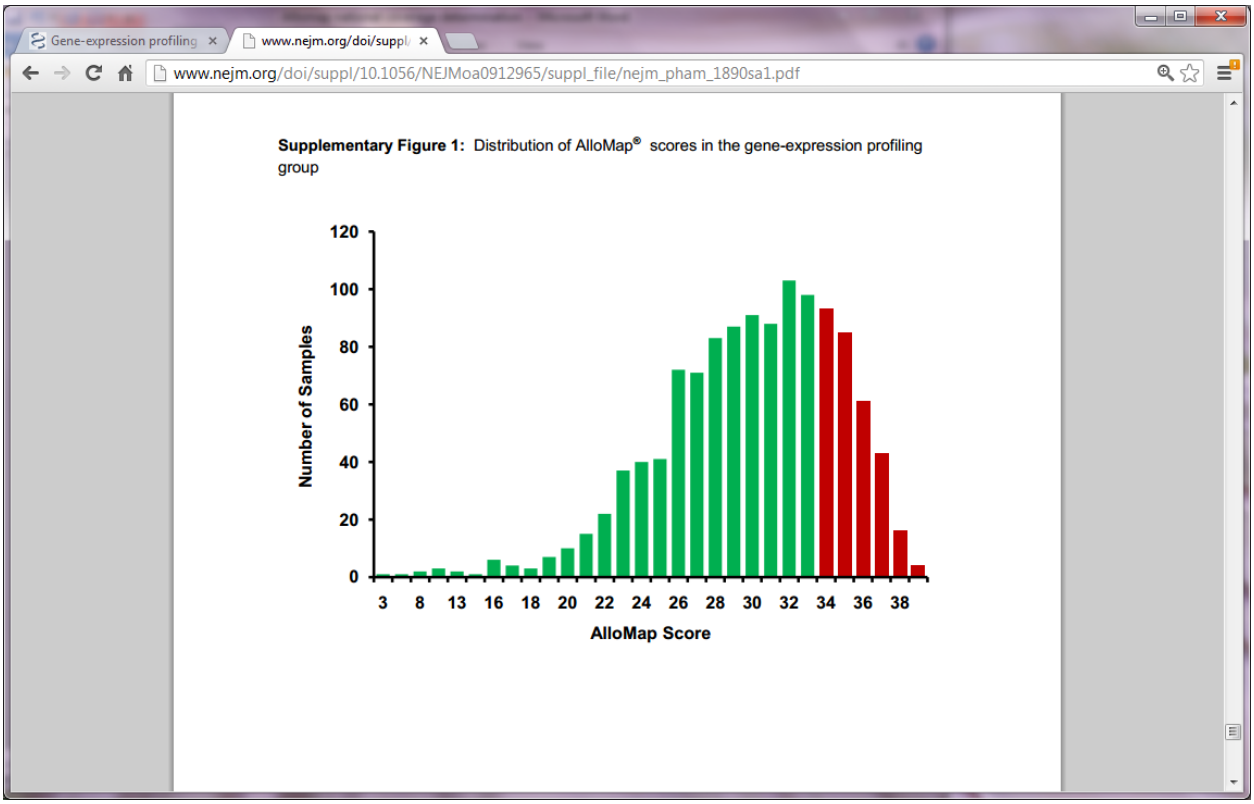
This data table was from the CARGO trial and indicates the range of AlloMap scores that were detected in samples in either the first 2-6 months or beyond 6 months after transplantation. The dynamic range of the test reported here is between 19 and 39 (based on a 0-40 scale), with 22.4% and 5.4% of patient samples being <19 in the 2-6 and >6 month groups respectively. In either time course (2-6 or >6 months) the PPV essentially is never above 5%. Thus, the test does not have a useful PPV. Curiously though, the NPV is never less than 97.9%. In the CARGO trial the authors used an AlloMap score of 30 which had a NPV of 98.7%. One could ask why a value of 37 wasn't used (NPV of 98.4%) or a 22 wasn't used (NPV of 100%). In fact one can take any AlloMap score of 19-39 and report a very good NPV value (between 98.3% and 100%) in the >6 month range. The only reason to slide up or down this scale would be to modulate the number of followup biopsies to

be performed – not to significantly improve the NPV or PPV of the test. In fact, in the IMAGE trial, the AlloMap cut off was changed from 30 to 34 for that exact reason. This is the rational given in the supplemental data of the Pham MX NEJM manuscript for that change:

> The initial protocol specified a gene-expression profiling score of 30 or higher to prompt a required endomyocardial biopsy. This threshold was selected based upon the initial findings from the CARGO study, showing that a score below 30 was associated with a negative predictive value of 99.6% for concurrent ISHLT Grade 3A (2R) or higher rejection. On November 7, 2005, a protocol amendment increased the threshold to 34 to minimize the number of biopsies needed in the gene-expression profiling group without compromising the assay performance. The decision to increase the threshold was based upon additional analyses from the CARGO study showing that the negative predictive value of the gene-expression profiling test remained robust (99.2%) at a higher threshold of 34 while reducing the number of positive tests from 50.8% to 22.3%.

Thus, the IMAGE trial collaborators were comfortable with a 0.4% change in NPV. This change significantly reduced the number of biopsies to be performed, improving the perceived usefulness of the test and was one of the key reported findings of the manuscript, as stated in the abstract: “Patients who were monitored with the use of gene-expression profiling underwent fewer biopsies per person-year of follow-up than did patients who were monitored with the use of endomyocardial biopsies (0.5 vs. 3.0, $P<0.001$).”

This second image from the supplemental data (supplementary figure 1) shows the distribution of AlloMap scores. There are two messages from this image. The first is that as you move the threshold for biopsy from 30 to 34, you greatly reduce the number of biopsies to perform. Secondly, the scores have a near normal distribution with a small left-sided tail. That is not an ideal distribution for a test that is supposed to discriminate between who should or should not receive a follow up biopsy.



An additional argument against the test is the actual sensitivity and specificity of the test. Although these were not reported in the manuscript, they could be inferred in several ways from the supplemental data (page 17). In the AlloMap group, there were 34 cases of treated rejection (grade 2R or 3R). Of these, 20 (59%) were treated due to overt heart failure, but were not identified by the AlloMap test. Seven individuals had both a positive AlloMap score and heart failure and a further 6 had a positive AlloMap score with concurrent biopsy showing rejection, but without over clinical symptoms. So the sensitivity of the test in identifying acute cellular rejection causing overt heart failure can be calculated as ((27-20)/27) or 0.26. The sensitivity of the test to identify a biopsy positive for grade 2R or 3R rejection is ((6+7)/34) or 0.38.

There were 1190 AlloMap tests performed, of which 302 (25%) were ≥34. Of these 302, 274 (91%) necessitated a biopsy. Of these 274, 265 were in asymptomatic patients. Of these 265 biopsies, there were 11 grade 2R or 3R results. So a positive AlloMap test, without accompanying symptoms, was predictive of rejection 4% (11/265) of the time. Additionally, of the 1190 AlloMap tests, 888 gave a value <34. In these individuals, although we do not know how many would have a positive biopsy as they were not performed, we know that 20 had overt clinical signs of rejection. Thus a negative AlloMap missed rejection a minimum of 2% (20/888) of the time and likely more often as biopsy-proved rejection is often noted without overt clinical symptoms. It is likely, but not proven, that 2R or 3R rejection was equal in those individuals with high or low AlloMap test results.

In summation of the AlloMap test statistics one can see how there is no discernment between any AlloMap test score between 19 and 39 and that values can be arbitrarily chosen to determine how many followup biopsies should be performed. Secondly, the test has a very poor sensitivity and the percent of patients identified by an AlloMap positive score as having rejection (4%) is essentially the same as the percent of patients with a negative AlloMap score who had overt clinical signs of rejection (2%).

It is my belief that a blood based assay to determine acute cellular and antibody-mediated rejection is a very good idea based on all of the limitations of the endomyocardial biopsy listed above. However, this specific test, the XDx AlloMap test does not perform adequately to warrant its use or for Medicare to pay for its use.

In a conversation with an XDx representative who was trying to convince me to use the test clinically at my institution, after going through the data with her, she agreed the test is essentially a placebo for clinicians. I whole-heartedly agree with that position.

Sincerely yours,

Marc Halushka MD, PhD
Department of Pathology
Johns Hopkins Medical Institutions

REFERENCES:

Stewart S, Winters GL, Fishbein MC, Tazelaar HD, Kobashigawa J, Abrams J, Andersen CB, Angelini A, Berry GJ, Burke MM, Demetris AJ, Hammond E, Itescu S, Marboe CC, McManus B, Reed EF, Reinsmoen NL, Rodriguez ER, Rose AG, Rose M, Suciu-Focia N, Zeevi A, Billingham ME. Revision of the 1990 working formulation for the standardization of nomenclature in the diagnosis of heart rejection. J Heart Lung Transplant. 2005 Nov;24(11):1710-20.

Horwitz PA, Tsai EJ, Putt ME, Gilmore JM, Lepore JJ, Parmacek MS, Kao AC, Desai SS, Goldberg LR, Brozena SC, Jessup ML, Epstein JA, Cappola TP. Detection of cardiac allograft rejection and response to immunosuppressive therapy with peripheral blood gene expression. Circulation. 2004 Dec 21;110(25):3815-21

Deng MC, Eisen HJ, Mehra MR, Billingham M, Marboe CC, Berry G, Kobashigawa J, Johnson FL, Starling RC, Murali S, Pauly DF, Baron H, Wohlgemuth JG, Woodward RN, Klingler TM, Walther D, Lal PG, Rosenberg S, Hunt S; CARGO Investigators. Noninvasive discrimination of rejection in cardiac allograft recipients using gene expression profiling. Am J Transplant. 2006 Jan;6(1):150-60.

Pham MX, Deng MC, Kfoury AG, Teuteberg JJ, Starling RC, Valantine H. Molecular testing for long-term rejection surveillance in heart transplant recipients: design of the Invasive Monitoring Attenuation Through Gene Expression (IMAGE) trial. J Heart Lung Transplant. 2007 Aug;26(8):808-14.

Pham MX, Teuteberg JJ, Kfoury AG, Starling RC, Deng MC, Cappola TP, Kao A, Anderson AS, Cotts WG, Ewald GA, Baran DA, Bogaev RC, Elashoff B, Baron H, Yee J, Valantine HA; IMAGE Study Group. Gene-expression profiling for rejection surveillance after cardiac transplantation. N Engl J Med. 2010 May 20;362(20):1890-900.

Revision of the 1990 Working Formulation for the Standardization of Nomenclature in the Diagnosis of Heart Rejection

Susan Stewart, FRCPath, Chair: Gayle L. Winters, MD, Michael C. Fishbein, MD, Henry D. Tazelaar, MD, Jon Kobashigawa, MD, Jacki Abrams, MD, Claus B. Andersen, MD, DMSc, Annalisa Angelini, MD, Gerald J. Berry, MD, Margaret M. Burke, FRCPath, Anthony J. Demetris, MD, Elizabeth Hammond, MD, Silviu Itescu, MD, Charles C. Marboe, MD, Bruce McManus, MD, PhD, Elaine F. Reed, PhD, Nancy L. Reinsmoen, PhD, E. Rene Rodriguez, MD, Alan G. Rose, MD, FRCPath, Marlene Rose, PhD, MRCPath, Nicole Suciu-Focia, PhD, Adriana Zeevi, PhD, and Margaret E. Billingham, FRCPath

In 1990, an international grading system for cardiac allograft biopsies was adopted by the International Society for Heart Transplantation. This system has served the heart transplant community well, facilitating communication between transplant centers, especially with regard to patient management and research. In 2004, under the direction of the International Society for Heart and Lung Transplantation (ISHLT), a multidisciplinary review of the cardiac biopsy grading system was undertaken to address challenges and inconsistencies in its use and to address recent advances in the knowledge of antibody-mediated rejection. This article summarizes the revised consensus classification for cardiac allograft rejection. In brief, the revised (R) categories of cellular rejection are as follows: Grade 0 R—no rejection (no change from 1990); Grade 1 R—mild rejection (1990 Grades 1A, 1B and 2); Grade 2 R—moderate rejection (1990 Grade 3A); and Grade 3 R—severe rejection (1990 Grades 3B and 4). Because the histologic sub-types of Quilty A and Quilty B have never been shown to have clinical significance, the "A" and "B" designations have been eliminated. Recommendations are also made for the histologic recognition and immunohistologic investigation of acute antibody-mediated rejection (AMR) with the expectation that greater standardization of the assessment of this controversial entity will clarify its clinical significance. Technical considerations in biopsy processing are also addressed. This consensus revision of the Working Formulation was approved by the ISHLT Board of Directors in December 2004. J Heart Lung Transplant 2005;24:1710–20.

Change is one thing, progress is another.
Change is scientific, progress is ethical.
Change is indubitable, progress is a matter of controversy.
Bertrand Russell
British philosopher
(1872–1970)

At the request of the International Society for Heart and Lung Transplantation (ISHLT), a standardized grading system for the pathologic diagnosis of rejection in cardiac biopsies was developed in 1990 to address the proliferation of diverse grading systems that occurred during the 1980s. The goal was to develop a uniform description and grading scheme for acute cardiac rejection, to improve communication between clinicians and investigators, to enable comparison of treatment regimens and outcomes between transplant centers, to facilitate multicenter clinical trials, and to promote further studies to determine the clinical significance of the various histologic patterns.[1] It was also intended to have a grading system that was easily learned, readily usable, reproducible, had defined clinical end-points, and could be modified as new information became available. The 1990 ISHLT grading system for cardiac biopsies was widely adopted and served the heart transplant community well for over a decade. However, several issues have arisen during this period requiring re-evaluation of the grading scheme.

First, it has become apparent that there were widespread inconsistencies in the use of the grading system as highlighted by multicenter therapeutic trials in which central pathology panel reviewers have been used for confirmation of endomyocardial biopsy diagnoses.[2,3] Major areas of diagnostic difficulty have included: Grade 1A vs Grade 2; Grade 1B vs Grade 3B; Grade 2 vs Grades 3A or 3B; Quilty B vs Grade 2 or 3A; and ischemic injury vs Grades 2 or 3A. Less common and less problematic areas of difficulty have included biopsy site(s) vs Grade 2 or 3A, Quilty B vs post-

From the International Society for Heart and Lung Transplantation, Addison, Texas.

Submitted March 16, 2005; revised March 16, 2005; accepted March 30, 2005.

Reprint requests: Susan Stewart, FRCPath, Department of Pathology, Papworth Hospital, Papworth Everard, Cambridge CB3 8RE, UK. Telephone: 44-1480-364-308. Fax: 44-1480-364-777. E-mail: susan.stewart@papworth.nhs.uk.

 1053-2498/05/$–see front matter. doi:10.1016/j.healun.2005.03.019

Table 1. ISHLT Standardized Cardiac Biopsy Grading: Acute Cellular Rejection[b]

2004		1990	
Grade 0 R[a]	No rejection	Grade 0	No rejection
Grade 1 R, mild	Interstitial and/or perivascular infiltrate with up to 1 focus of myocyte damage	Grade 1, mild A—Focal	Focal perivascular and/or interstitial infiltrate without myocyte damage
		B—Diffuse	Diffuse infiltrate without myocyte damage
		Grade 2 moderate (focal)	One focus of infiltrate with associated myocyte damage
Grade 2 R, moderate	Two or more foci of infiltrate with associated myocyte damage	Grade 3, moderate A—Focal	Multifocal infiltrate with myocyte damage
Grade 3 R, severe	Diffuse infiltrate with multifocal myocyte damage ± edema, ± hemorrhage ± vasculitis	B—Diffuse	Diffuse infiltrate with myocyte damage
		Grade 4, severe	Diffuse, polymorphous infiltrate with extensive myocyte damage ± edema, ± hemorrhage + vasculitis

[a]Where "R" denotes revised grade to avoid confusion with 1990 scheme.
[b]The presence or absence of acute antibody-mediated rejection (AMR) may be recorded as AMR 0 or AMR 1, as required (see Table 3).

transplant lymphoproliferative disorder (PTLD) and infection or PTLD vs rejection.

When the 1990 grading system was proposed, the clinical importance of Grade 2 (focal moderate rejection) histology was unknown and, therefore, a separate rejection grade was assigned to allow studies to determine the clinical significance of this histologic pattern. The proposal was made at that time to meet again and review the data regarding the clinical correlations of the grades and to modify the system as necessary. It should also be noted that the 1990 grading system was defined in biopsies from patients generally receiving triple-drug therapy (steroids, cyclosporine, azathioprine) for immunosuppression. Since that time, immunosuppressive regimens have changed, the incidence of rejection has changed, and it is possible that the histology of rejection may also have changed.

The advances in the understanding of transplant rejection and new therapeutic options to prevent and/or treat rejection have warranted re-examination of the grading system. An attempt was made in 1994–1995 to fine-tune the 1990 grading system and clarify those areas that had caused difficulty in interpretation, including Grade 2 acute rejection.[4] This revision drew mixed responses and was never officially adopted or published. The grading system was again discussed at the Sixth Banff Conference on Allograft Pathology in 2001, where a working group exchanged ideas and experience in using the 1990 grading system and recommended a review and update of the grading system, including the need to establish clear criteria for the pathologic diagnosis of humoral rejection.[5] In 2004, again under the direction of the ISHLT, a multidisciplinary review of the cardiac biopsy grading system was undertaken with task forces examining the areas of histopathology/cellular rejection, humoral (antibody-mediated) rejection, clinical issues and future research. In addition, comments solicited from the ISHLT membership at large were taken into account, which mainly concerned

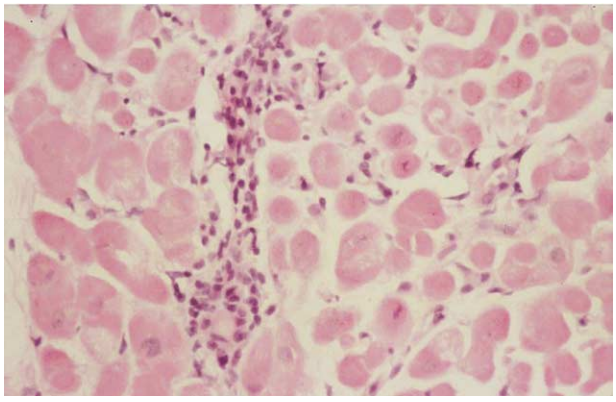


Figure 1. Myocardial biopsy showing acute cellular rejection with an inflammatory infiltrate composed of mainly lymphocytes in a perivascular distribution and not extending into interstitium or damaging myocytes. Hematoxylin and Eosin. (H&E)

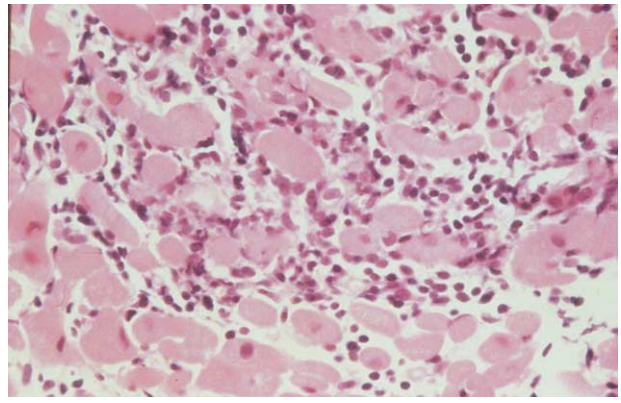


Figure 2. Myocyte damage characterized by encroachment of mononuclear cells at the perimeter of myocytes resulting in irregular, scalloped borders and distorting the cellular architecture. Several myocytes are surrounded by infiltrating cells. (H&E).

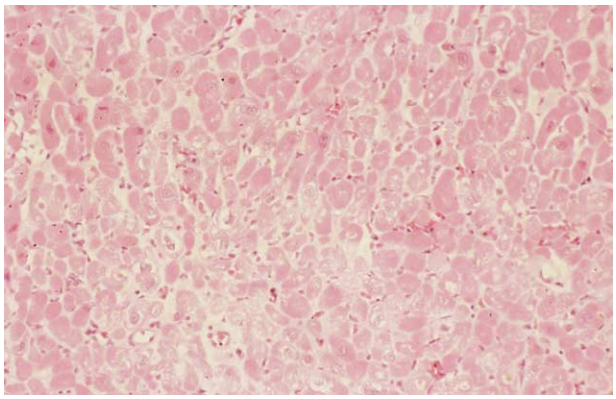


Figure 3. Grade 0 R: Normal endomyocardial biopsy showing no evidence of cellular infiltration. (H&E).

Grade 2 cellular rejection and humoral rejection. Consensus was reached and presented at the 24th Annual ISHLT scientific meeting. This study reports the consensus findings as a revision of the previous Working Formulation, which was approved by the ISHLT board in December 2004.

HISTOLOGIC DIAGNOSIS AND GRADING OF ACUTE CARDIAC ALLOGRAFT REJECTION

Biopsy-proven acute rejection on surveillance endomyocardial biopsies appears to be decreasing, due at least in part to improved immunosuppressive therapy.[6] In addition, there has been a shift in clinical response to some grades of rejection. In the middle to late 1980s, most (but not all) transplant centers treated any biopsy with myocyte injury (1990 ISHLT Grade 2 and higher) with some form of augmented immunosuppression, regardless of the clinical presentation. Several studies in the early to mid-1990s showed that lower grades of rejection resolve without treatment in a majority of cases.[7–14] Biopsies with 1990 ISHLT Grade 1, Grade 2 and even some sub-sets of Grade 3 rejection progress to high-grade rejection on the next biopsy in only 15% to 20% of cases. At the other end of the spectrum, Grades 3B and 4 are uniformly treated aggressively. Therefore, the consensus was to modify the 1990 ISHLT grading system as shown in Table 1. In brief:

- 1990 ISHLT Grades 1A, 1B and 2 would be combined into a new, revised 2004 ISHLT Grade 1 R.
- 1990 ISHLT Grade 3A would become 2004 ISHLT Grade 2 R; and
- 1990 ISHLT Grades 3B and 4 would become 2004 ISHLT Grade 3 R.

In addition, the Histopathology Task Force recommended that further characterization of the nature of the inflammatory infiltrate and definition of myocyte damage would be helpful in reducing inconsistencies in the application of the grading system (*vide infra*).

Inflammatory Infiltrate

Acute cellular rejection is characterized by an inflammatory infiltrate predominantly comprised of lympho-

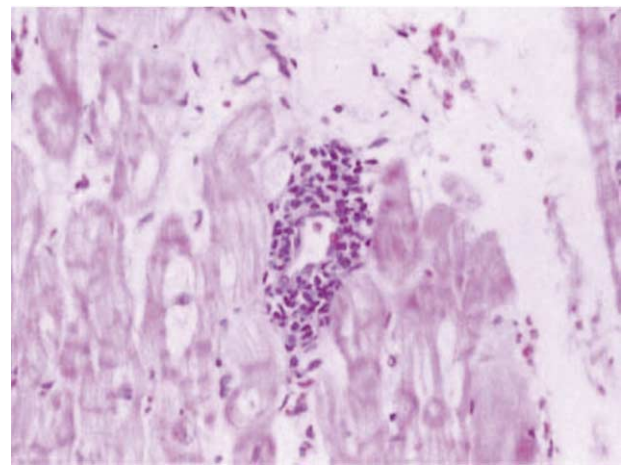


Figure 5. Grade 1 R: Higher power view of focal, perivascular mononuclear cell infiltrate without myocyte encroachment or damage. Previously Grade 1A. (H&E).

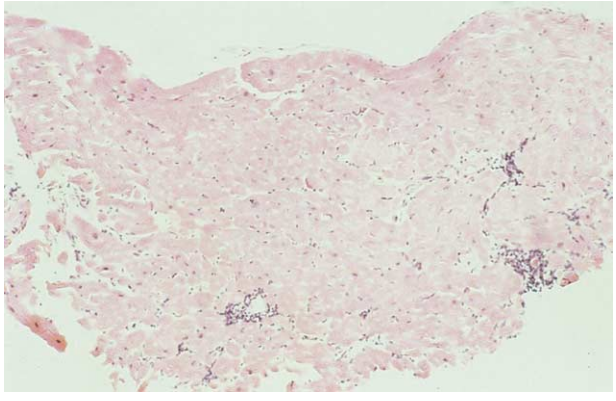


Figure 4. Grade 1 R: Low power view of endomyocardial biopsy showing three focal, perivascular infiltrates without myocyte damage. Previously Grade 1A (H&E).

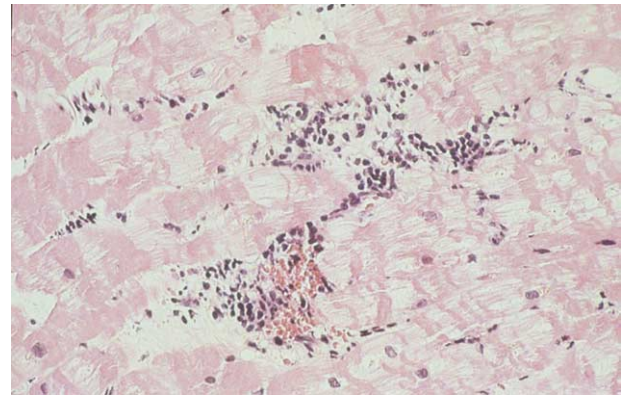


Figure 6. Grade 1 R: Both perivascular and interstitial infiltrates are present but without definite evidence of myocyte damage. Previously Grade 1A (H&E).

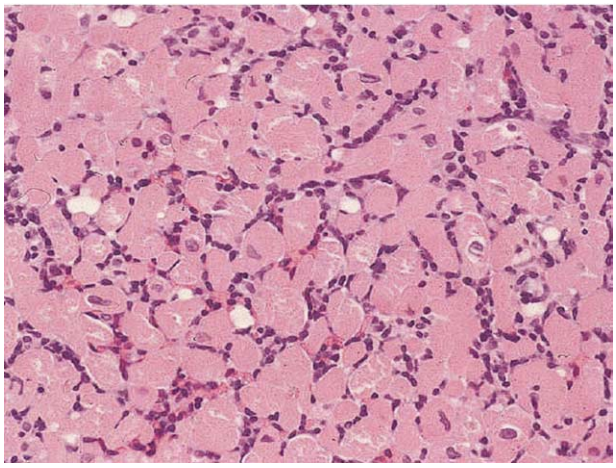


Figure 7. Grade 1 R: Diffuse mononuclear cell infiltrate with an interstitial pattern of lymphocytes between and around myocytes without associated myocyte damage. Previously Grade 1B. (H&E).

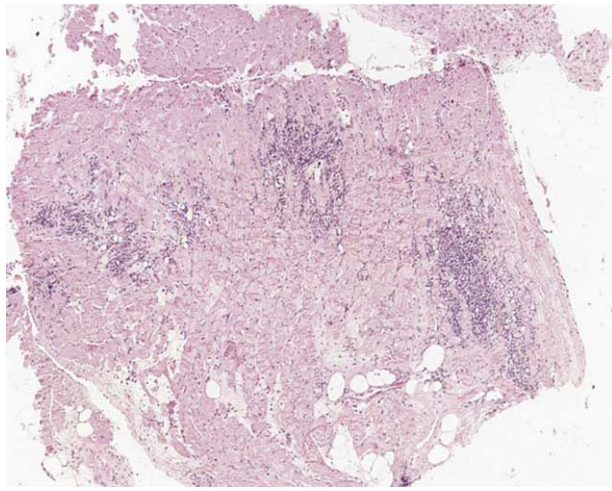


Figure 9. Grade 2 R: Low power view showing three foci of damaging mononuclear cell infiltrate with normal myocardium intervening. Previously Grade 3A. (H&E).

cytes, as well as macrophages and occasional eosinophils (Figure 1). The presence of neutrophils (except in the most severe form of rejection) should raise the question of an alternative process, such as healing ischemic injury, antibody-mediated (humoral) rejection or infection. Plasma cells are also not typically present in acute cellular rejection and suggest a Quilty lesion, healing ischemic injury (often in response to allograft coronary disease) or a lymphoproliferative disorder (plasmacytoid lymphocytes).

Myocyte Damage

Damage or injury to the myocardium, originally termed "myocyte necrosis," is an important but sometimes difficult feature to identify. Although readily distinguishable, cell death may be a feature of the most severe forms of rejection; myocyte damage in milder rejection is often characterized by myocytolysis and no contraction band or coagulation necrosis. Features of myocytolysis include clearing of the sarcoplasm and nuclei, with nuclear enlargement and occasionally prominent nucleoli. The presence of myocyte injury is frequently accompanied by encroachment of inflammatory cells at the perimeter of myocytes, resulting in irregular or scalloped myocyte borders, their partial or whole replacement, or distortion of the normal myocardial architecture (Figure 2). These features are often better appreciated by the examination of multiple levels of sectioning. It should also be noted that myocytolysis can be seen in both early and late ischemic injury.

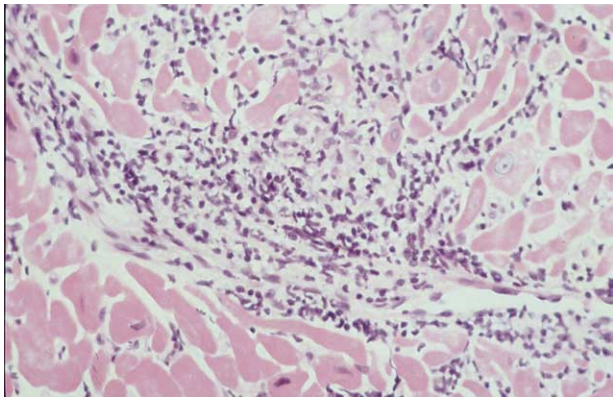


Figure 8. Grade 1 R: High power view of a mononuclear infiltrate extending from a perivascular position into adjacent myocardium with damage to myocytes and distortion of architecture. This is a single focus in the biopsy series and therefore is included in the revised mild grade of acute rejection, previously described as Grade 2. (H&E).

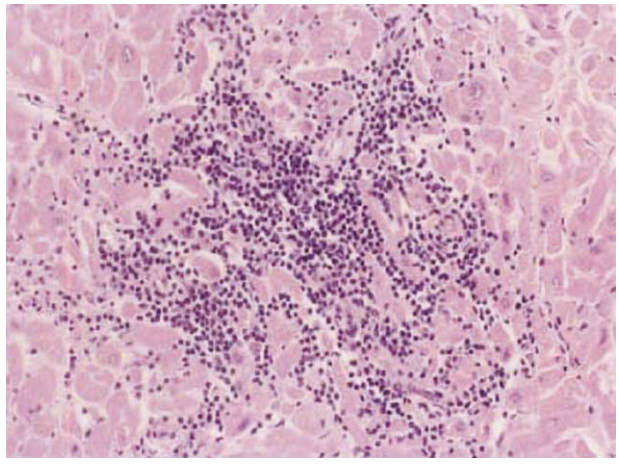


Figure 10. Grade 2 R: Higher power view of one focus of figure 9 damaging infiltrate with myocyte damage and architectural distortion (a "space occupying lesion"). (H&E).

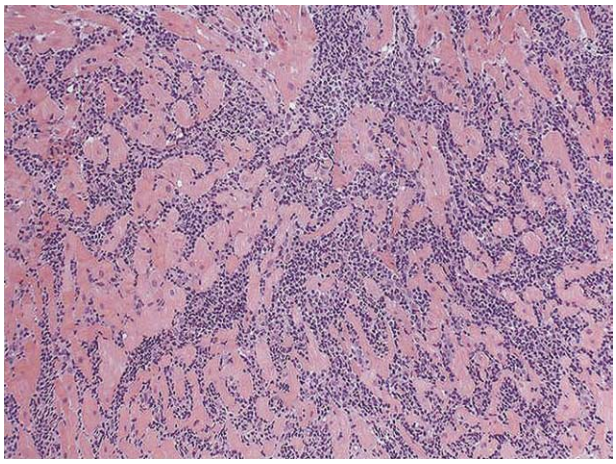


Figure 11. Grade 3 R: Diffuse damaging infiltrates with encroachment of myocytes and disruption of normal architecture. This contrasts with the non-damaging infiltrates of figure 7. Prevously Grade 3B. (H&E).

Grade 0 R (no acute cellular rejection)

In Grade 0 R there is no evidence of mononuclear (lymphocytes/macrophages) inflammation or myocyte damage (Figure 3).

Grade 1 R (mild, low-grade, acute cellular rejection)

Mild or low-grade rejection may manifest in one of two ways: (1) Perivascular and/or interstitial mononuclear cells (lymphocytes/histiocytes) are present. In general, these cells respect myocyte borders, do not encroach on adjacent myocytes, and do not distort the normal architecture (Figures 4, 5, 6 and 7). (2) One focus of mononuclear cells with associated myocyte damage may be present (Figures 2 and 8).

Table 2. Nonrejection Biopsy Findings

2004	1990
Ischemic injury	Ischemic injury
Early—up to 6 weeks post-transplant	A = up to 3 weeks post-transplant
Late—related to allograft coronary disease	B = late ischemia
Quilty effect	Quilty effect
	A = no myocyte encroachment
	B = with myocyte encroachment
Infection	Infection
Lymphoproliferative disorder	Lymphoproliferative disorder

Grade 2 R (moderate, intermediate-grade, acute cellular rejection)

In Grade 2 R two or more foci of mononuclear cells (lymphocytes/macrophages) with associated myocyte damage are present. Eosinophils may be present. The foci may be distributed in one or more than one biopsy fragment. Intervening areas of uninvolved myocardium are present between the foci of rejection (Figures 9 and 10). Low-grade (Grade 1R) rejection can be present in other biopsy pieces.

Grade 3 R (severe, high-grade, acute cellular rejection)

A diffuse inflammatory process, either predominantly lymphocytes and macrophages or a polymorphous infiltrate, is present, involving multiple biopsy fragments (Figures 11 and 12). In most cases, the majority of biopsy fragments are involved, although the intensity of the infiltrate may vary between pieces. Multiple areas of associated myocyte damage are present. In the most severe forms of cellular (and humoral) rejection,

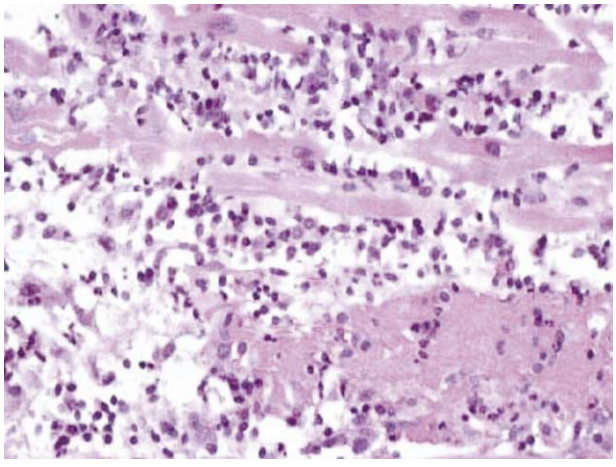


Figure 12. Grade 3 R: Severe acute rejection with widespread myocyte damage and some necrosis. The diffuse infiltrate includes polymorphs as well as lymphocytes, macrophages and plasma cells. Previously Grade 4. (H&E).

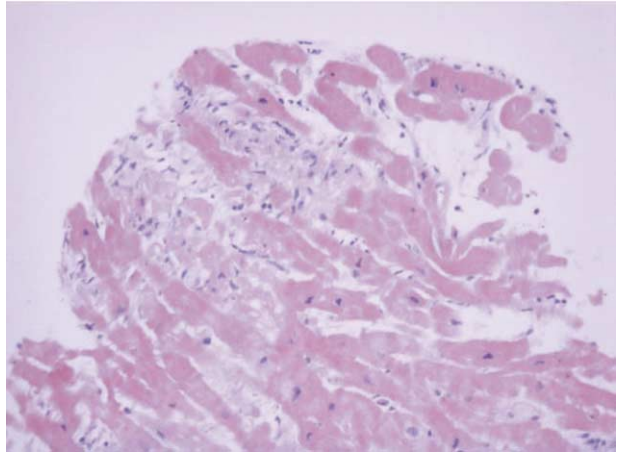


Figure 13. Peritransplant injury showing a focus of ischemic injury with myocytolysis and vacuolization. Note the relative lack of infiltrating inflammatory cells compared with acute cellular rejection. Macrophages are present. (H&E).

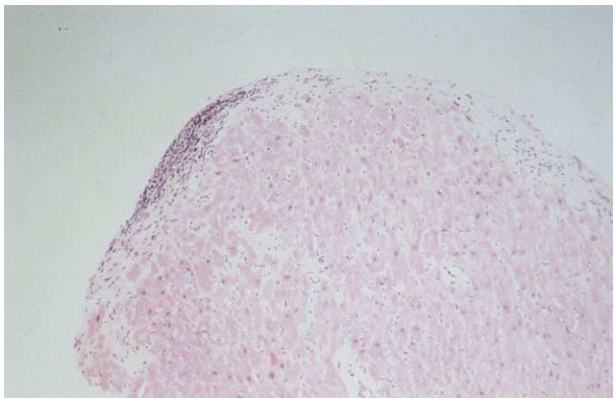


Figure 14. Low power view of non-encroaching endocardial infiltrate or Quilty lesion with normal underlying myocardium. (H&E).

edema, interstitial hemorrhage and vasculitis may be present.

NON-REJECTION BIOPSY FINDINGS

Peri-operative Ischemic Injury

Early (peri-operative) ischemic injury arises in the peri-operative period during the obligatory ischemic time that accompanies procurement and implantation of a donor heart (Table 2).[15] Such injury may be exacerbated by prolonged hypotension due to poor graft function, hemorrhage during the peri-operative period, and the effects of prolonged high-dose inotrope therapy. Ischemic injury is characterized initially by contraction band necrosis or coagulative myocyte necrosis, often with myocyte vacuolization and fat necrosis, and frequently extends to the endocardial surface. As healing ensues, biopsies may contain mixed inflammatory infiltrates, including

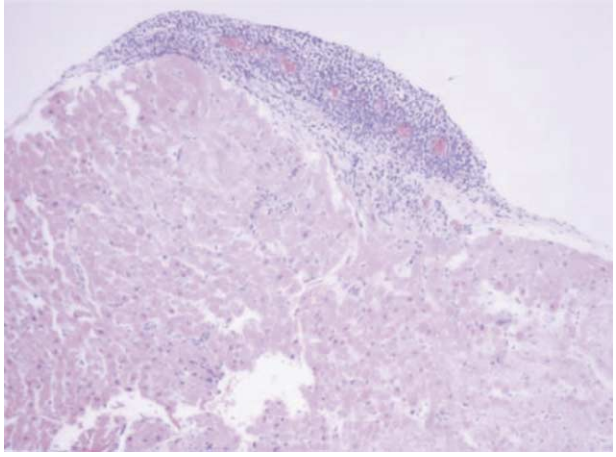


Figure 15. Higher power view of another area of the same biopsy as figure 14, showing some superficial encroachment of the endocardial lesion into underlying myocardium. Note the prominent vascularity of this endocardial infiltrate which can be a very useful feature for distinguishing tangentially cut infiltrates from foci of acute cellular rejection. (H&E).

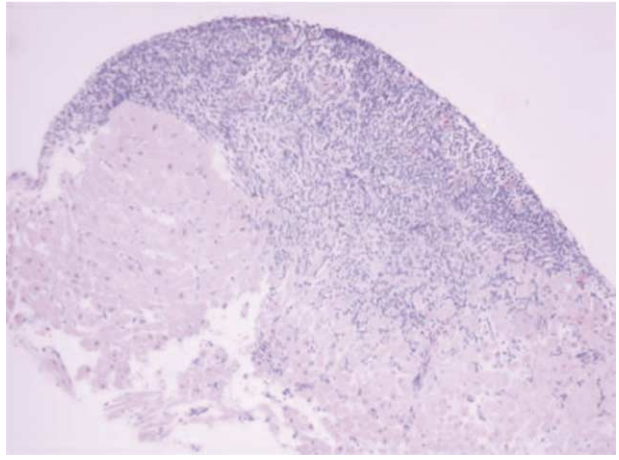


Figure 16. A deeper section of the biopsy in figure 15 showing much greater encroachment into myocardium and less vascularity. (H&E).

neutrophils as well as lymphocytes, macrophages and eosinophils, and it is at this point that confusion with acute rejection may occur. Ischemic injury, especially in its healing phase, is a common biopsy finding in the early post-transplant period (up to 6 weeks) and must be differentiated from acute rejection. In acute rejection, the inflammatory infiltrate frequently is proportionally greater than the degree of myocyte damage, whereas, in ischemic foci, it is usually the reverse (Figure 13). Peri-transplant injury with neutrophils may show overlapping features with antibody-mediated (humoral) rejection (*vide infra*).

Late Ischemic Injury (related to allograft coronary disease)

Assessing the arterial changes of allograft coronary disease in endomyocardial biopsy specimens is usu-

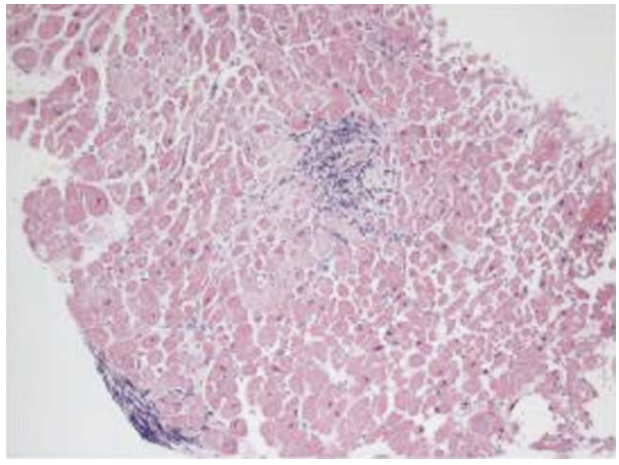


Figure 17. Endomyocardial biopsy showing a small endocardial infiltrate and focus of deeper intramyocardial cellular infiltration which raises the possibility of acute cellular rejection until deeper sections are examined. (H&E).

ally precluded by the lack of vessels large enough to permit such an evaluation. However, the ability to detect secondary myocardial changes, such as myocyte vacuolization and microinfarcts, may be helpful in determining the etiology of late cardiac failure.[16] In addition, the diagnosis of late ischemic injury may be helpful in determining the etiology of cardiac failure in transplant recipients. It may be especially helpful in ruling out other potentially treatable etiologies that are part of the differential diagnosis, such as acute rejection or PTLD.

Quilty Effect

Nodular endocardial infiltrates, or Quilty effect, occur in approximately 10% to 20% of post-transplant endomyocardial biopsies.[17,18] The infiltrates may be confined to the endocardium (1990 ISHLT Quilty A) or may extend into the underlying myocardium where associated myocyte damage may be present (1990 ISHLT Quilty B) (Figures 14, 15 and 16). The histologic sub-typing of Quilty A and Quilty B has never been shown to have any clinical significance and there is agreement that separating Quilty A from B has no clinical value.[19] The designations "A" and "B" have therefore been eliminated and the lesion is referred to simply as the Quilty effect.

The relationship of Quilty effect to acute rejection, if any, remains unknown. Traditionally, this lesion has been considered distinct from acute rejection, requiring no treatment with intensified immunosuppression. Differentiation of Quilty effect from acute rejection is not usually a problem when the former is confined to the endocardium. However, when it extends into the underlying myocardium, a tangential cut through the biopsy may not show a connection between the myocardial lesion and the endocardial lesion, making differentiation from acute rejection more difficult.[20] Cutting additional deeper sections may resolve this dilemma in some cases by demonstrating extension to the endocardium (Figures 17 and 18). In the absence of an endocardial extension, the density of the infiltrate, presence of B lymphocytes and plasma cells, background fibrosis and prominent vascularity favor a diagnosis of Quilty effect.

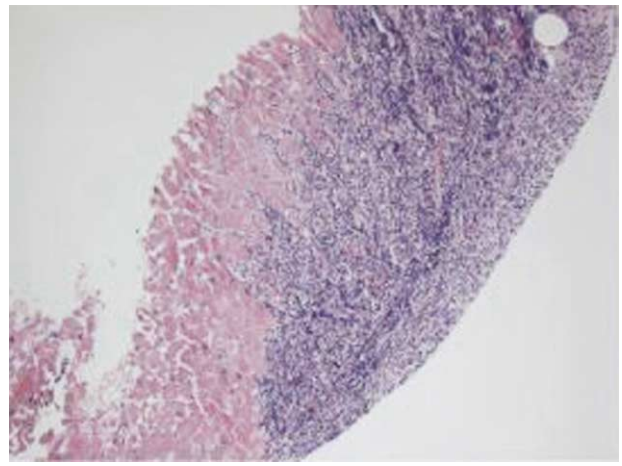


Figure 18. Deeper section of figure 17 clearly shows extension of the surface endocardial infiltrate into myocardium confirming the correct diagnosis of Quilty lesion rather than acute cellular rejection. (H&E).

Immunohistochemical staining of the infiltrate for B and T cells may be helpful in this regard.

Infection and PTLDs

Infection and PTLDs remain important causes of post-transplant morbidity and mortality, but are relatively rare in post-transplant cardiac biopsies. Notable among these are cytomegalovirus (CMV) and toxoplasmosis, both of which may be accompanied by lymphocyte-predominant infiltrates, which may be misinterpreted as acute cellular rejection, leading to inappropriate augmentation of immunosuppression. More specifically, targeted immunosuppression and improved prophylaxis protocols, especially for CMV, have decreased the incidence of some infections. Recognition of the relationship between immunosuppression and post-transplant neoplasms, especially PTLD, has favored less aggressive immunosuppression protocols. Although infection and PTLD are less controversial than other post-transplant biopsy interpretations, they require continued awareness and vigilance.

ACUTE ANTIBODY-MEDIATED (HUMORAL) REJECTION

Acute humoral rejection is recognized as a clinical entity in the grafted heart (Table 3). It remains controversial, however, with a highly varied incidence be-

Table 3. ISHLT Recommendations for Acute Antibody-Mediated Rejection (AMR)

	2004	1990
AMR 0	Negative for acute antibody-mediated rejection No histologic or immunopathologic features of AMR	Humoral rejection (positive immunoflourescence, vasculitis or severe edema in absence of cellular infiltrate) recorded as additional required information
AMR 1	Positive for AMR Histologic features of AMR Positive immunofluorescence or immunoperoxidase staining for AMR (positive CD68, C4d)	

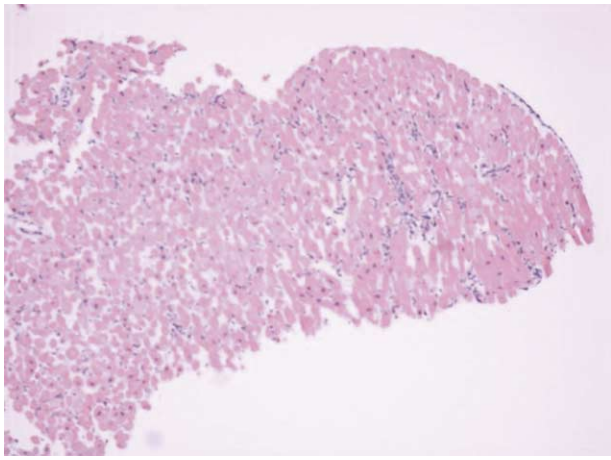


Figure 19. Antibody mediated rejection (AMR 1). Low power view of endomyocardial biopsy with scattered cellular infiltrates and intervening normal tissue. (H&E).

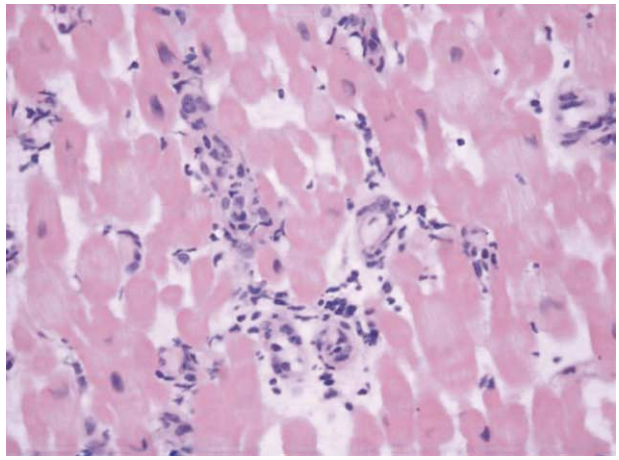


Figure 20. AMR 1. Higher power view shows that the cellular infiltrates are within vessels and include polymorphs. Endothelial cell swelling is present. The increased cellularity seen at low power is due to the presence of these intravascular cells and not perivascular inflammation. Compare with figures 1 and 5. (H&E).

tween different centers and no consensus has yet been reached on its recognition and diagnosis either histopathologically or immunologically.[21–25] The 2004 ISHLT meeting reviewed evidence from the immunopathology and clinical task forces and felt able to suggest diagnostic criteria in specific circumstances so that further assessment of this entity could be encouraged. In the 1990 Working Formulation, there was an option to record immunofluorescence studies for those centers that sought to pursue these on biopsy specimens in the first 6 weeks after transplantation.[1] Similarly, in utilizing the 2004 classification, pathologists can follow the guidance if they intend to investigate the possibility of antibody-mediated rejection as a cause of cardiac dysfunction. A separate companion study from the Immunopathology Task Force is available with a detailed discussion of antibody-mediated rejection. A summary of recommendations is provided here to allow incorporation, as required, into the revised Working Formulation.

Acute antibody-mediated rejection is associated with worse graft survival and is observed in allosensitized patients, including those with previous transplantation, transfusion or pregnancy and previous ventricular assist device use. The incidence may be up to 15% in the first year post-transplantation and the clinical presentation has no pathognomonic features. Pathologically, it can be recognized by myocardial capillary injury with endothelial-cell swelling and intravascular macrophage accumulation (Figures 19, 20 and 21). Interstitial edema and hemorrhage can be present together with neutrophils in and around capillaries. Intravascular thrombi and myocyte necrosis without cellular infiltration can also be identified.[21,22] When these features are seen in the presence of unexplained cardiac dysfunction, typically early onset of hemodynamic compromise and myocardial dysfunction, it is proposed that immunostaining can be performed by immunofluorescence or immunohistochemistry as follows:

- Immunoglobulin (IgG, IgM and/or IgA) plus complement deposition (C3d, C4d and/or C1q) in capillaries by immunofluorescence on frozen sections (Figures 22 and 23); and/or

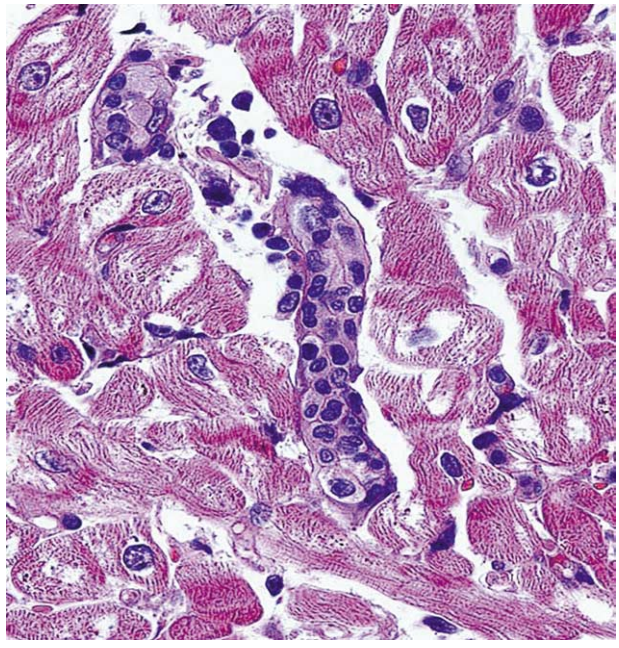


Figure 21. AMR 1. High power view confirms the intravascular location of the cells which have the appearance of macrophages and illustrates the endothelial cell swelling. (H&E).

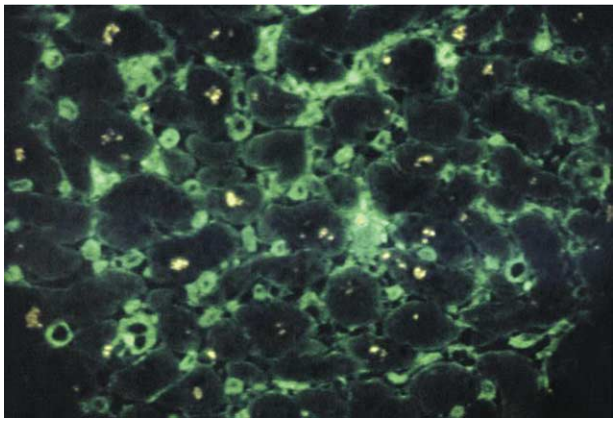


Figure 22. AMR 1. Immunofluorescence positivity for IgG clearly shown in and around capillaries.

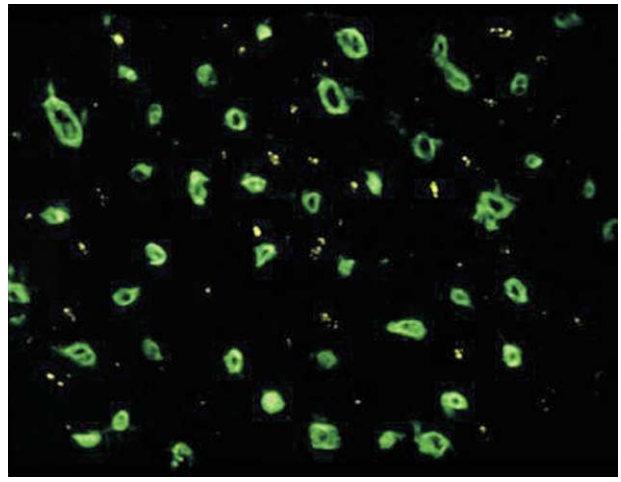


Figure 23. AMR 1. Immunofluorescence positivity for C4d in capillaries with characteristic "doughnut" appearance.

- CD68 staining of macrophages within capillaries (CD31- or CD34-positive) by immunohistochemistry (Figure 24); and
- C4 d staining of capillaries by paraffin immunohistochemistry (Figure 25).

It is recommended that patients with hemodynamic compromise undergo assessment for circulating antibodies.

The consensus meeting recommended that screening should not be advocated at this time, but every endomyocardial biopsy should undergo critical histologic evaluation for features suggestive of antibody-mediated rejection. If such features (as just detailed) are not seen, the biopsy should be designated negative for antibody-mediated rejection, or AMR 0. If features suggestive of antibody-mediated rejection are seen, the diagnosis of acute antibody-mediated rejection should be confirmed by immunohistochemistry, either immunofluorescence or immunoperoxidase, using antibodies directed against CD68, CD31 and C4d, and a serum should be drawn and tested for donor-specific antibody.[26,27] If these markers are positive, a positive diagnosis for AMR should be made (AMR 1). Patients who have several episodes of documented acute antibody-mediated rejection should be followed on future biopsies with at least one of these immunohistochemical methods and monitored for the production of donor-specific antibodies. It is also recognized that acute cellular and antibody-mediated rejection can co-exist, but further studies will be required to delineate these.

This recommended approach to the diagnosis of acute antibody-mediated rejection—if there is either a clinical indication or a research need—should encourage clinicians, histopathologists and immunologists to work together and clarify its existence, frequency and clinical significance.[28]

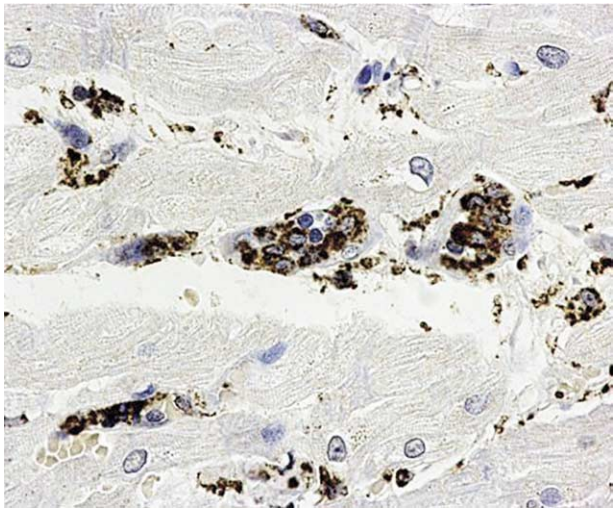


Figure 24. AMR 1. Immunoperoxidase staining is strongly positive for CD68, confirming the intravascular cells to be macrophages.

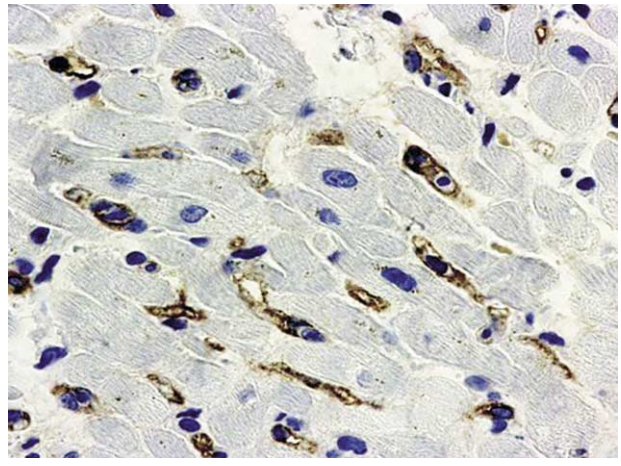


Figure 25. AMR 1. Immunoperoxidase staining is strongly positive for C4d in capillaries allowing a diagnosis of AMR to be made in the appropriate context. (see text).

Table 4. Technical Considerations

Minimum number of biopsy samples = 3
Number of hematoxylin and eosin slides = 3
Number of levels = 3
Routine special stains required = None

TECHNICAL CONSIDERATIONS

Due to the potential for sampling error in diagnosing acute rejection, multiple myocardial biopsy samples should be obtained from different right ventricle sites (Table 4). Samples should not be divided once procured in order to obtain the required number of pieces because this practice results in less representative sampling. Although the original ISHLT grading system required 4 samples of myocardium, the trend has been to accept 3 evaluable samples as the absolute minimum for interpretation. Therefore, a minimum of 3, and preferably 4 (or more), evaluable pieces of myocardium are now recommended for grading acute cellular rejection. An evaluable piece of myocardium contains at least 50% myocardium, excluding previous biopsy site, scar, adipose tissue or blood clot, which may comprise the remainder of the piece. Hematoxylin and eosin staining of at least 3 levels through the tissue samples are recommended. Additional spare slides may be saved unstained if additional studies are needed. Special stains are not routinely required and have been eliminated by many centers as the incidence of rejection has decreased. A trichrome stain may be helpful in selected cases for assessing myocyte damage and fibrosis, such as in the early post-operative period.

CONCLUSIONS

It is the intention of this consensus group that this revision of the grading system addresses and clarifies concerns that have developed in the 15 years since the adoption of the 1990 grading system. The plan is to supplement this revision with an educational program for pathologists and clinicians. As was the case for the 1990 grading system, the 2004 grading system will now be required for all ISHLT-sponsored meetings and publications.

There has been tremendous advancement in technology since the 1990 grading system was instituted, including many molecular techniques. Many of these advances have been used successfully in the research setting to further our knowledge of pathologic processes. The challenge will be to decide the appropriate time and choice of technique(s) to incorporate into routine clinical practice. For the ISHLT grading system to remain the standard worldwide, it must remain the lowest common denominator so that every transplant center has the technical ability and financial resources to incorporate any proposed changes. We must make sure, going forward, that we retain the universality of the grading system because this has always been a major component of its success. The consensus meeting task forces strongly urge the ISHLT to periodically review the grading system as immunosuppressive regimens evolve and as additional clinical and molecular monitors of cardiac function, coronary vasculopathy and immune responsiveness are developed and used in the management of heart transplant recipients.

REFERENCES

1. Billingham ME, Cary NRB, Hammond EH, et al. A Working Formulation for the Standardization of Nomenclature in the Diagnosis of Heart and Lung Rejection: Heart Rejection Study Group. J Heart Transplant 1990;9:587–93.
2. Winters GL, McManus BM, for the Rapamycin Cardiac Rejection Treatment Trial Pathologists. Consistencies and controversies in the application of the ISHLT Working Formulation for cardiac transplant biopsy specimens. J Heart Lung Transplant 1996;15:728–35.
3. Lindop GBM, Burke MM, Ogston S, et al. Inter-observer variability in grading acute rejection in endomyocardial biopsies. J Heart Lung Transplant 2003;22(suppl):S33.
4. Winters GL, Marboe CC, Billingham ME. The ISHLT grading system for cardiac transplant biopsies: clarification and commentary. J Heart Lung Transplant 1998;17:754–60.
5. Rodriguez ER. The pathology of heart transplant biopsy specimens: revisiting the 1990 ISHLT Working Formulation. J Heart Lung Transplant 2003;22:3–15.
6. Winters GL. Heart transplantation. In: Thiene G, Pessine AC, editors. Advances in cardiovascular medicine. Padua, Italy: University of Padua; 2002:145–63.
7. Laufer G, Laczkovics A, Wollenek G, et al. The progression of mild acute cardiac rejection evaluated by risk factor analysis: the impact of maintenance steroids and serum creatinine. Transplantation 1991;51:184–9.
8. Yeoh TK, Frist WH, Eastburn TE, Atkinson J. Clinical significance of mild rejection of the cardiac allograft. Circulation 1992;86 (5 suppl):II-267-71.
9. Lloveras JJ, Escourrou G, Delisle MB, et al. Evolution of untreated mild rejection in heart transplant recipients. J Heart Lung Transplant 1992;11:751–6.
10. Gleeson MP, Kobashigawa JA, Stevenson LW, et al. The natural history of focal moderate rejection in orthotopic heart transplant recipients. JAMA 1994;43:483A.
11. Rizeq MN, Masek MA, Billingham ME. Acute rejection: significance of elapsed time after transplantation. J Heart Lung Transplant 1994;13:862–8.
12. Winters GL, Loh E, Schoen FJ. Natural history of focal moderate cardiac allograft rejection: is treatment warranted? Circulation 1995;91:1975–80.
13. Milano A, Calories ALP, Livi U, et al. Evaluation of focal moderate (ISHLT Grade 2) rejection of the cardiac allograft. J Heart Lung Transplant 1996;15:456–60.
14. Brunner-LaRocca HP, Sutsch G, Schneider J, et al. Natural course of moderate cardiac allograft rejections (Interna-

tional Society of Heart and Lung Transplantation Grade 2) early and late after transplantation. Circulation 1996;94:1334–8.
15. Fyfe B, Loh E, Winters GL, et al. Heart transplantation-associated perioperative ischemic myocardial injury: morphological features and clinical significance. Circulation 1996;93:1133–40.
16. Winters GL, Schoen FJ. Graft arteriosclerosis-induced myocardial pathology in heart transplant recipients: predictive value of endomyocardial biopsy. J Heart Lung Transplant 1997;16:985–91.
17. Kottke-Marchant K, Ratliff NB. Endomyocardial lymphocytic infiltrates in cardiac transplant recipients. Arch Pathol Lab Med 1989;113:690–8.
18. Radio SJ, McManus BM, Winters GL, et al. Preferential endocardial residence of B-cells in the "Quilty effect" of human heart allografts: immunohistochemical distinction from rejection. Mod Pathol 1991;4:654.
19. Joshi A, Masek MA, Brown BW, et al. Quilty revisited: a 10-year perspective. Hum Pathol 1995;26:547–57.
20. Fishbein MC, Bell G, Lones MA, et al. Grade 2 cellular heart rejection: does it exist? J Heart Lung Transplant 1994;13:1051–7.
21. Hammond EH, Yowell RL, Nunoda S, et al. Vascular (humoral) rejection in heart transplantation: pathologic observations and clinical implications. J Heart Transplant 1989;8:430–43.
22. Lones MA, Lawrence SC, Alfredo T, Deborah HRN, Miller JM, Fishbein MC. Clinical pathologic features of humoral rejection in cardiac allografts: a study of 81 consecutive patients. J Heart Lung Transplant 1995;14:151–62.
23. Michaels PJ, Espejo ML, Kobashigawa J, et al. Humoral rejection in cardiac transplantation: risk factors, hemodynamic consequences and relationship to transplant coronary artery disease. J Heart Lung Transplant 2003;22:58–69.
24. Loh E, Bergin JD, Couper GS, Mudge GH Jr. Role of panel-reactive antibody cross-reactivity in predicting survival after orthotopic heart transplantation. J Heart Lung Transplant 1994;13:194–201.
25. Taylor DO, Yowell RL, Khoury AG, Hammond EH, Renlund DG. Allograft coronary artery disease: clinical correlations with circulating anti-HLA antibodies and the immunohistopathologic pattern of vascular rejection. J Heart Lung Transplant 2000;19:518–21.
26. Chantranuwat C, Qiao JH, Kobashigawa J, Hong L, Shintaku P, Fishbein MC. Immunohistochemical staining for C4d on paraffin embedded tissue in cardiac allograft endomyocardial biopsies: comparison to frozen tissue immunofluorescence. Appl Immunohistochem Mol Morphol 2004;12:166–71.
27. Baldwin WM III, Samaniego-Picota M, Kasper EK, et al. Complement deposition in early cardiac transplant biopsies is associated with ischemic injury and subsequent rejection episodes. Transplantation 1999;68:894–900.
28. Takemoto S, Zeevi A, Feng S, et al. A national conference to assess antibody mediated rejection in solid organ transplantation. Am J Transplant 2004;4:1033–41.

APPENDIX: PARTICIPANTS BY TASK FORCE

Chair of Consensus Meeting: Susan Stewart, FRCPath.

Histopathology

Chair: Gayle L. Winters, MD. *Participants:* Jacki Abrams, MD; Claus B. Andersen, MD, DMSc; Annalisa Angelini, MD; Gerald J. Berry, MD; Margaret M. Burke, FRCPath; Anthony J. Demetris, MD; Michael C. Fishbein, MD; Charles C. Marboe, MD; Bruce M. McManus, MD, PhD; Alan G. Rose, MD, FRCPath; Henry D. Tazelaar, MD.

Immunopathology

Chair: Michael C. Fishbein, MD. *Participants:* Elizabeth Hammond, MD; Silviu Itescu, MD; Elaine F Reed, PhD; Nancy L. Reinsmoen PhD; E. Rene Rodriguez, MD; Marlene Rose, PhD, MRCPath; Nicole Suciu-Foca, PhD; Adriana Zeevi, PhD.

Clinical Heart Transplantation

Chair: Jon Kobashigawa, MD. *Participants:* Manfred Hummel, MD, PhD; Sharon Hunt, MD; Anne Keogh, MD; James K. Kirklin, MD; Mandeep Mehra, MD; Leslie W. Miller, MD; Paul Josef Mohasci, MD; Jayan Parameshwar, MRCP; Branislav Radovancevic, MD; Heather J. Ross, MD; Randall Starling, MD.

Research

Chair: Anthony J. Demetris, MD. *Participants:* Bruce M. McManus, MD, PhD; E. Rene Rodriguez, MD; Gayle L. Winters, MD.

Detection of Cardiac Allograft Rejection and Response to Immunosuppressive Therapy With Peripheral Blood Gene Expression

Phillip A. Horwitz, Emily J. Tsai, Mary E. Putt, Joan M. Gilmore, John J. Lepore, Michael S. Parmacek, Andrew C. Kao, Shashank S. Desai, Lee R. Goldberg, Susan C. Brozena, Mariell L. Jessup, Jonathan A. Epstein and Thomas P. Cappola

Circulation. 2004;110:3815-3821; originally published online December 6, 2004;
doi: 10.1161/01.CIR.0000150539.72783.BF
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 0009-7322. Online ISSN: 1524-4539

Detection of Cardiac Allograft Rejection and Response to Immunosuppressive Therapy With Peripheral Blood Gene Expression

Phillip A. Horwitz, MD; Emily J. Tsai, MD; Mary E. Putt, ScD, PhD; Joan M. Gilmore, BS; John J. Lepore, MD; Michael S. Parmacek, MD; Andrew C. Kao, MD; Shashank S. Desai, MD; Lee R. Goldberg, MD, MPH; Susan C. Brozena, MD; Mariell L. Jessup, MD; Jonathan A. Epstein, MD; Thomas P. Cappola, MD, ScM

Background—Assessment of gene expression in peripheral blood may provide a noninvasive screening test for allograft rejection. We hypothesized that changes in peripheral blood expression profiles would correlate with biopsy-proven rejection and would resolve after treatment of rejection episodes.

Methods and Results—We performed a case-control study nested within a cohort of 189 cardiac transplant patients who had blood samples obtained during endomyocardial biopsy (EMB). Using Affymetrix HU133A microarrays, we analyzed whole-blood expression profiles from 3 groups: (1) control samples with negative EMB (n=7); (2) samples obtained during rejection (at least International Society for Heart and Lung Transplantation grade 3A; n=7); and (3) samples obtained after rejection, after treatment and normalization of the EMB (n=7). We identified 91 transcripts differentially expressed in rejection compared with control (false discovery rate <0.10). In postrejection samples, 98% of transcripts returned toward control levels, displaying an intermediate expression profile for patients with treated rejection ($P<0.0001$). Cluster analysis of the 40 transcripts with >25% change in expression levels during rejection demonstrated good discrimination between control and rejection samples and verified the intermediate expression profile of postrejection samples. Quantitative real-time polymerase chain reaction confirmed significant differential expression for the predictive markers *CFLAR* and *SOD2* (UniGene ID No. 355724 and No. 384944).

Conclusions—These data demonstrate that peripheral blood expression profiles correlate with biopsy-proven allograft rejection. Intermediate expression profiles of treated rejection suggest persistent immune activation despite normalization of the EMB. If validated in larger studies, expression profiling may prove to be a more sensitive screening test for allograft rejection than EMB. (***Circulation*. 2004;110:3815-3821.**)

Key Words: immune system ■ transplantation ■ rejection ■ genes ■ diagnosis

Detection of allograft rejection is a major clinical concern in the care of heart transplant recipients. The optimal approach is to detect rejection before the onset of cardiac dysfunction and to treat it aggressively with augmented immunosuppression. It is equally important to reduce immunosuppression in patients who do not have rejection to minimize drug toxicity. The current standard to screen for rejection is the detection of inflammatory infiltrates on serial endomyocardial biopsy (EMB)[1,2]; however, EMB is an invasive procedure limited by patient discomfort, risk of complications, and cost.[3] These barriers prevent frequent monitoring for rejection and limit optimal titration of immunosuppressive therapy.

Rejection is a complex immune response that involves T-cell recognition of alloantigens in the cardiac allograft, costimulatory signals, elaboration of effector molecules by activated T cells, and an inflammatory response within the graft.[4–10] Activation and recruitment of circulating leukocytes to the allograft is an essential part of this process, which makes peripheral blood monitoring of the immune response an attractive method for the noninvasive detection of rejection. The purpose of the present study was to test the hypothesis that gene-expression profiles obtained from peripheral blood correlate with histological cardiac allograft rejection on serial EMBs. Our findings raise the possibility that peripheral blood gene-expression profiles

Received June 29, 2004; revision received September 28, 2004; accepted October 15, 2004.

From the Department of Internal Medicine, Division of Cardiovascular Diseases, University of Iowa–Roy J. and Lucille A. Carver College of Medicine (P.A.H.), Iowa City, Iowa; Department of Medicine, Division of Cardiovascular Medicine, University of Pennsylvania School of Medicine (E.J.T., J.M.G., J.J.L., M.S.P., A.C.K., S.S.D., L.R.G., S.C.B., M.L.J., J.A.E., T.P.C.), Philadelphia, Pa; and Division of Biostatistics, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania (M.E.P.), Philadelphia, Pa.

Correspondence to Thomas P. Cappola, MD, ScM, Heart Failure and Transplantation, Division of Cardiovascular Medicine, 6 Penn Tower, 3400 Spruce St, Philadelphia, PA 19104. E-mail thomas.cappola@uphs.upenn.edu

Circulation **is available at http://www.circulationaha.org** **DOI: 10.1161/01.CIR.0000150539.72783.BF**

could serve as a noninvasive method to screen for cardiac allograft rejection.

Methods

Patient Population

We prospectively collected 409 blood samples from 189 consecutive cardiac transplant patients referred for routine surveillance EMB at the Hospital of the University of Pennsylvania between March and July 2002. All subjects gave written informed consent, and the University of Pennsylvania Institutional Review Board approved the study protocol.

Sample Collection

Blood samples were obtained from a central venous sheath immediately before EMB and were collected in RNA preservation solution (PAXgene Blood RNA Tubes, Qiagen Inc) for immediate RNA stabilization and storage at −80°C. EMB specimens were assessed by a cardiac pathologist at the University of Pennsylvania, and rejection grade was determined with the International Society for Heart and Lung Transplantation (ISHLT) grading system.[1] This system categorizes biopsies into several grades (0, 1A, 1B, 2, 3A, 3B, and 4) based on the extent of lymphocyte infiltration, myocyte necrosis, and presence or absence of hemorrhage. Augmented immunosuppression is indicated for ISHLT grade 3A or higher rejection.

Study Design

We performed a nested case-control study of peripheral blood gene expression within our cohort of biopsy patients. Case patients ("rejection") were chosen on the basis of the presence of rejection severe enough to mandate augmented immunosuppression according to our clinical protocols (IHSLT grade 3A or higher).[11] Control patients were selected on the basis of the absence of clinically significant rejection (ISHLT grade 1A or lower). To minimize clinical confounders, both rejection and control samples were obtained from patients who had no clinical evidence of active infection or other acute illness at the time of biopsy and whose recent clinical status had been stable at least 1 week before their scheduled biopsy. All case and control patients were treated with standard immunosuppression with corticosteroids, antimetabolites, calcineurin inhibitors, and/or sirolimus.

In addition, we selected follow-up blood specimens from the rejection patients after treatment with augmented immunosuppression and resolution of rejection to grade 2 or lower on repeat EMB. This group of postrejection specimens ("postrejection") allowed us to analyze changes in gene-expression profile over time in the same patients during and after resolution of clinically significant rejection.

Microarray Sample Preparation and Hybridization

Control (n=7), rejection (n=7), and postrejection (n=7) samples were selected as described above and purified with a commercial nucleic acid isolation column (PAXgene Blood RNA Column, Qiagen Inc). Total RNA samples were analyzed by Agilent bioanalzyer and OD_{260}/OD_{280} ratio for RNA quality and quantification. Individual complementary DNAs (cDNAs) were prepared from each RNA isolate with reverse transcriptase [Superscript II primed by a poly (T) oligomer/T7 promoter]. Each cDNA was subsequently used as a template to make biotin-labeled cRNA with an in vitro transcription reaction, which resulted in a single cRNA for each blood sample. Each cRNA was hybridized with an individual Affymetrix HU133A oligonucleotide array, which was subsequently processed and scanned according to the manufacturer's instructions. All arrays (n=21) were hybridized on the same day by a single technician to avoid variability in hybridization conditions. Each array quantifies the expression of 22 215 transcripts (including full-length mRNA sequences and expressed sequence tags) derived from build 133 of the UniGene database (available at www.affymetrix.com). Data were saved as raw image files and converted into probe-set data (.cel files) with Microarray Suite (MAS 5.0).

Microarray Analysis

There are several methods to convert Affymetrix probe-set data into normalized measures of gene expression, including software provided by the manufacturer (MAS5), model-based methods (dCHIP), and robust multiarray analysis (RMA).[12,13] We chose RMA on the basis of its superiority in the analysis of small data sets.[14,15] Software for RMA is available (www.bioconductor.org) for use in the R 1.70 package for statistical computing (www.r-project.org).[16]

Differentially Expressed Genes in Rejection Compared With Control Samples

To determine candidate markers of rejection, we applied 3 criteria to the normalized data. First, data were filtered to include genes present above background on at least 1 array. Second, significance analysis of microarrays (SAM; available at http://www-stat.stanford.edu/≈tibs/SAM/) was used to correct for multiple comparisons and to select candidate markers of rejection using genes that were differentially expressed with an estimated overall false-discovery rate <0.10.[17] Third, we required at least a 25% change in expression between rejection and control samples for a transcript to be of interest. The identities of differentially expressed genes were determined with annotation databases (available at www.netaffx.com) or via BLAST searches (http://www.ncbi.nih.gov/BLAST) of the corresponding expressed sequence tags.

Response to Treatment

To determine whether our candidate markers of rejection responded to immunosuppressive therapy, we analyzed expression data for these transcripts in postrejection samples. If our candidate genes were markers of rejection, we hypothesized that genes that were overexpressed (underexpressed) in rejection versus control should also be overexpressed (underexpressed) in rejection versus postrejection. Alternatively, if our candidate genes were identified owing to confounding factors (eg, differences in age between the rejection and control subjects), then we would not expect the pattern of differential expression to be recapitulated in the paired rejection versus postrejection comparison. First, we determined the proportion of the 91 candidate genes in which the direction of the fold change for rejection versus control was concordant with the direction of the fold change for rejection versus postrejection. Individual candidate genes were scored as concordant if fold changes that were greater than (less than) 1.0 for rejection versus control were also greater than (less than) 1.0 for rejection versus postrejection. Second, we estimated the probability of selecting a set of 91 candidates by chance that had the observed degree of concordance or a higher degree of concordance by randomly selected sets of 91 genes from the rejection versus postrejection array data. Thus, we randomly selected 91 genes, determined whether each gene was concordant, and computed the total number of concordant genes in the randomly selected group. We repeated this process 10 000 times, and used the repeated samples to determine our probability value, ie, the probability of a chance occurrence of the observed or better concordance.

Cluster Analysis

The capacity of our candidate markers to distinguish control, rejection, and postrejection samples was assessed by hierarchical clustering. Clusters were constructed with average linkage clustering and Pearson correlation coefficients as a measure of similarity with Cluster software and displayed with Treeview software (available at http://rana.lbl.gov).[17]

Validation

Quantitative real-time polymerase chain reaction (qRT-PCR) was performed to validate changes in selected genes. Because there was insufficient RNA after the microarray studies to validate these data from all of the original samples, validation was performed with mRNA harvested from additional samples from the original biopsy cohort by the same selection criteria. RNA isolates were treated with DNAse to remove any contaminating genomic DNA and were subsequently converted to cDNA with an in vitro transcription reaction. cDNAs were used as templates for Taqman qRT-PCR with

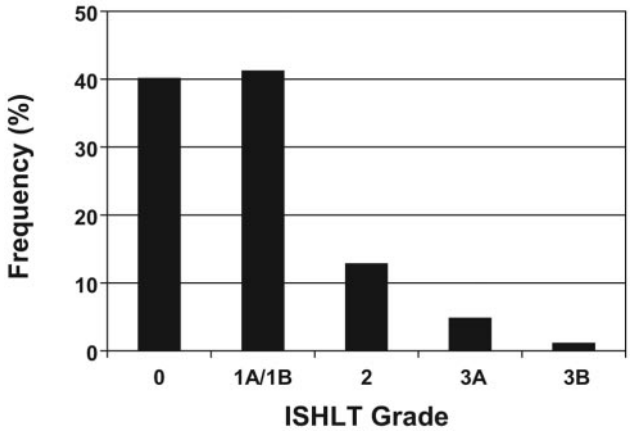


Figure 1. Distribution of rejection by ISHLT biopsy grade in cohort of transplant recipients (189 transplant recipients; 409 biopsies).

ABI Assays-on-Demand on an ABI Prism 7900 sequence detection system. The specific assays used were Hs00153439_m1 (CFLAR), Hs00167309_m1 (SOD2), and Hs99999905_m1 (GAPDH). All samples were run in triplicate, and GADPH was used as an internal control to normalize transcript abundance. Triplicates were averaged to calculate an expression value for each sample. Data were compared among control, rejection, and postrejection samples by the Wilcoxon rank-sum test, with $P<0.05$ indicating statistical significance.

Results

Patient Characteristics

The frequency of rejection was low in the present study population (Figure 1). Of 409 EMB samples, 81% showed minimal or no evidence of allograft rejection (ISHLT grades 0, 1A, or 1B), and 6% showed clinically significant rejection (grade 3A or higher) that required increases in the immunosuppression regimen. The characteristics of patients chosen for study are outlined in Table 1. All control samples had grade 0 rejection on biopsy, and all rejection samples were obtained from patients with rejection graded 3A or higher. The postrejection samples were obtained a median of 55 days after rejection was first detected.

Microarray Analysis

Candidate Markers of Rejection

Of the 22 215 transcripts on each array, 10 826 (49%) were expressed at levels higher than background in at least 1 of the 21 samples. Of these, 91 gene products were differentially expressed in rejection compared with control (Figure 2, red) with a false-discovery rate <0.10 after SAM analysis. Seven genes were overexpressed and 84 genes were underexpressed in rejection. These genes were regarded as candidate markers for high-grade rejection. Overall, there was good reproducibility in gene expression in these candidates. The average coefficient of variation within each group (control or rejection) was 4%; however, reproducibility was different for each gene, ranging from a minimum coefficient of variation of 1% to a maximum of 11%.

We assessed changes in our candidate markers after treatment of rejection by measuring expression levels in follow-up samples from the same patients. As shown in Figure 2 (blue), expression of nearly all of the candidate markers moved closer to a fold change of 1 after immunosuppressive therapy, which indicates a return toward levels in control. This finding

TABLE 1. Patient Characteristics

	Control (n=7)	Rejection (n=7)	Postrejection* (n=7)
Biopsy grade			
0	7	0	3
1A	0	0	2
1B	0	0	2
2	0	0	0
3A	0	5	0
3B	0	2	0
Age, y, median (range)	61 (54–67)	45 (28–66)	45 (28–66)
Gender			
Female	1	2	2
Male	6	5	5
Type of immunosuppression†			
Double therapy	2	1	0
Triple therapy	5	6	7
Days after transplantation, median (range)	326 (8–1259)	491 (7–1865)	...
Days of augmented immunosuppression, median (range)	...	...	55 (14–76)

*Rejection and postrejection represent the same patients during and after treatment of grade 3A or higher rejection with augmented immunosuppression.

†Double therapy indicates mycophenolate, azathioprine, or sirolimus plus cyclosporine or tacrolimus; triple therapy indicates double therapy plus corticosteroids.

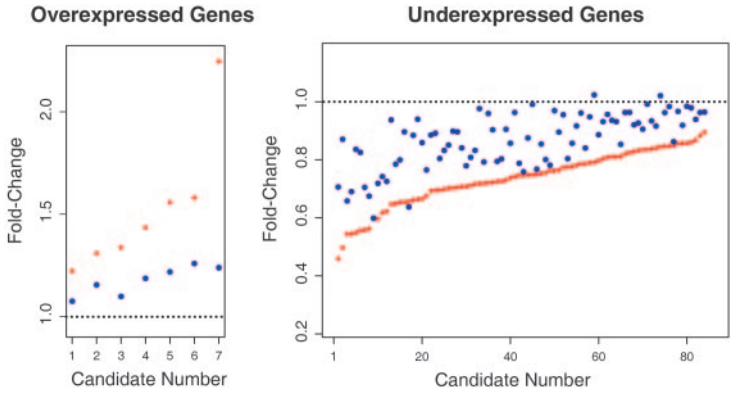


Figure 2. Differential gene expression in peripheral blood specimens from patients with biopsy-proven transplant rejection (n=7) and controls without rejection (n=7). As shown (red), 7 genes were overexpressed and 84 genes were underexpressed in rejection. After treatment and resolution of rejection on follow-up EMB, follow-up microarray analysis in these same patients (n=7) demonstrated that expression levels returned toward level in control (blue; $P<0.0001$ by resampling).

is consistent with the response to therapy noted on EMB; however, expression in the postrejection samples did not fully normalize to a fold change of 1, which suggests that treated rejection has an intermediate expression profile between control and rejection. By randomly resampling gene-expression data, we estimated the probability of finding a set of 91 genes that by chance showed differential expression in rejection with concordant changes after rejection. Only 1 in 10 000 randomly selected sets of 91 genes showed this pattern ($P=0.0001$); therefore, it is extremely unlikely that the observed intermediate expression profile occurred owing to chance. These findings suggest that we have identified an expression profile that correlates with active rejection in these patients.

Cluster Analysis

We used hierarchical clustering as an additional method to characterize the ability of our candidate markers to distinguish control, rejection, and postrejection samples. Hierarchical clustering is a computational method that groups experimental samples according to similarity in patterns of gene expression across a large number of genes.[18] We selected 40 transcripts that showed at least a 25% change in expression between control and rejection and performed cluster analysis on this panel of genes. As shown in Figure 3, samples clustered into 2 main branches, with complete partitioning of control and rejection samples into separate branches. Postrejection samples were present in both the control and rejection branches of the dendrogram, consistent with an intermediate expression profile for treated rejection.

Gene Function

The identities of our 40 candidate markers of rejection included 30 unique transcripts (Table 2). The majority of these are involved in the following cellular pathways: (1) transcription or translation, (2) cell-cycle regulation, (3) tumorigenesis/tumor suppression, (4) immune response, (5) apoptosis, and (6) intracellular signaling. Also included in Table 2 are a number of expressed sequence tags of unknown function. Several transcripts are represented by multiple probe sets on the HU133A array. These replicate probe sets showed consistent changes during rejection that resolved at postrejection biopsy time points (Figure 3). The marker with the largest number of internal replicates was the gene *CASP8* and FADD-like apoptosis regulator (CFLAR), an inhibitor of apoptosis that is downregulated in rejection.

Quantitative PCR

We verified transcriptional changes using qRT-PCR for 2 genes: *CFLAR* and superoxide dismutase 2 (*SOD2*). Consistent with the microarray analysis, both genes were significantly downregulated during rejection, with a mean fold change of 0.76 ± 0.06 ($P=0.01$) for *CFLAR* and a mean fold change of 0.74 ± 0.09 ($P=0.02$) for *SOD2*, as shown in Figure 4. Thus, peripheral blood gene-expression changes observed by microarray profiling were confirmed in comparisons of rejection and control samples. In postrejection samples, *CFLAR* expression trended back toward control levels, with a fold change closer to 1.0, but *SOD2* did not. The partial return toward baseline for *CFLAR* and the lack of return for *SOD2* likely reflect persistent partial activation of circulating leukocytes in these samples, which were taken at variable times after histological resolution of rejection.

Discussion

This study demonstrates the principle that peripheral blood gene expression correlates with cardiac allograft rejection detected on EMB. We identified 40 transcripts that are altered in acute cellular rejection and returned toward normal in response to augmented immunosuppression. Moreover, we observed in 2 separate analyses that treated rejection has an intermediate expression profile, which suggests persistent immune activation despite resolution of rejection on biopsy. These findings raise the possibility that expression profiling may prove to be a more sensitive screening test for rejection than EMB.

Previous investigators have used molecular markers to develop better screening tests for cardiac allograft rejection. For example, expression of immune stimulatory and activation markers (CD40, CD27, TIRC7), cytokines (interferon-γ, interleukin [IL]-2, IL-4, IL-6, IL-8), and cytotoxic T-cell effector molecules (perforin, granzyme B, FasL) are elevated in biopsy samples of rejecting myocardium.[4–10] These markers could be used to enhance the sensitivity of biopsy-detected rejection, but they do not eliminate the need for invasive procedures. Other groups have correlated levels of circulating markers, such as cytokine or cytokine mRNA levels, with cardiac allograft rejection in an effort to develop noninvasive screening tests.[19,20] In particular, Morgun et al[21] performed quantitative PCR analysis on peripheral mononuclear cell candidate transcripts and found correlations between EMB results and candidate mRNA expression levels.

TABLE 2. Candidate Expression Markers of Cardiac Allograft Rejection

Gene (Gene Symbol)	Protein Type/Cellular Pathway	Fold Change (Rejection vs Control)	Fold Change (Postrejection vs Control)	Probe-Set ID*	UniGene ID†
Ubiquinol-cytochrome C reductase binding protein (*UQCRB*)	Oxidative respiration	2.25	1.3	205849_s_at	Hs0.131255
Basic transcription factor 3 (*BTF3*)	RNA translation	1.57	1.24	208517_x_at 211939_x_at	Hs0.446567
Suppression of tumorigenicity 13 (*ST13*)	Tumor suppressor	1.43	1.19	207040_s_at	Hs0.377199
Cullin 4A (*CUL4A*)	Cell cycle/DNA replication	1.34	1.1	201423_s_at	Hs0.270788
Telomeric repeat binding factor 2, interacting protein (*TERF2IP*)	Transcription factor	1.31	1.15	201174_s_at	Hs0.274428
Arrestin, beta 2 (*ARRB2*)	Intracellular signaling	0.75	0.79	203388_at	Hs0.435811
EST		0.75	0.99	207365_x_at	Hs0.435123
EST		0.74	0.96	207730_x_at	Hs0.406701
EST		0.74	0.86	205781_at	Hs0.164410
Aminopeptidase puromycin sensitive (*NPEPPS*)	Proteinase	0.73	0.91	201454_s_at	Hs0.293007
Phosphatidylinositol glycan, class B (*PIGB*)	Cell surface protein	0.73	0.8	205452_at	Hs0.259326
Adenomatosis polyposis coli (*APC*)	Tumor suppressor	0.72	0.9	216933_x_at	Hs0.75081
B-cell CLL/lymphoma 7A (*BCL7A*)	Cell cycle/DNA replication	0.72	0.98	210679_x_at	Hs0.371758
Endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 4 (*EDG4*)	Cell cycle/DNA replication	0.72	0.81	206722_s_at 206723_s_at	Hs0.122575
Interleukin 17 receptor (*IL17R*)	Interleukin receptor	0.72	0.79	205707_at	Hs0.129751
Placental growth factor (*PGF*)	Hormone/angiogenesis factor	0.72	0.96	215179_x_at	Hs0.252820
EST		0.7	0.85	220712_at	Hs0.493129
EST		0.7	0.9	215558_at	Hs0.485406
EST		0.7	0.9	220071_x_at	Hs0.14347
Nuclear factor of activated T-cells 5, tonicity-responsive (*NFAT5*)	Transcription factor	0.7	0.83	208003_s_at	Hs0.86998
EST		0.69	0.89	221205_at	
EST		0.69	0.89	215978_x_at	Hs0.447720
Baculoviral IAP repeat-containing 1 (*BIRC1*)	Apoptosis	0.68	0.76	204861_s_at	Hs0.79019
Leukocyte immunoglobulin-like receptor, subfamily B, member 3 (*LILRB3*)	Leukocyte receptor	0.68	0.81	210784_x_at 211135_x_at	Hs0.306230
EST		0.66	0.94	209703_x_at	Hs0.288771
Transmembrane 6 superfamily member 2 (*TM6SF2*)	Cell surface protein	0.66	0.88	210598_at	Hs0.367829
EST		0.65	0.9	215375_x_at	Hs0.438377
EST		0.65	0.94	215029_at	Hs0.293563
CASP8 and FADD-like apoptosis regulator (*CFLAR*)	Apoptosis	0.59	0.73	211862_x_at 210564_x_at 208485_x_at 211317_s_at 214486_x_at	Hs0.355724
Superoxide dismutase 2, mitochondrial (*SOD2*)	Oxidative stress	0.56	0.83	221477_s_at	Hs0.384944
EST		0.55	0.84	216109_at	Hs0.435249
Solute carrier family 16, member 3 (*SLC16A3*)	Membrane transport	0.54	0.66	202856_s_at	Hs0.386678
Stearoyl-CoA desaturase 4 (*SCD4*)	Fatty acid metabolism	0.5	0.87	220232_at	Hs0.379191

EST indicates expressed sequence tag.
*Probe-Set ID indicates the corresponding probe set on the Affymetrix HU 133A microarray (available at http://www.affymetrix.com).
†UniGene ID (available at http://www.ncbi.nlm.nih.gov/UniGene).

These studies support the hypothesis that peripheral blood gene expression may reflect organ-level rejection, but they are limited by the short list of candidate markers tested with PCR-based technologies.

In contrast to previously employed methods, microarray technology offers the possibility of simultaneously screening thousands of potential candidate genes in an unbiased fashion. These advantages allow for the identification of gene-expression profiles that may be much more sensitive and specific than any one candidate marker, as has been shown in previous studies of hematologic malignancies and renal transplantation.[22–25] The relatively small fold changes in gene expression observed in the present study (<2.5-fold) support the hypothesis that an aggregate marker composed of multi-

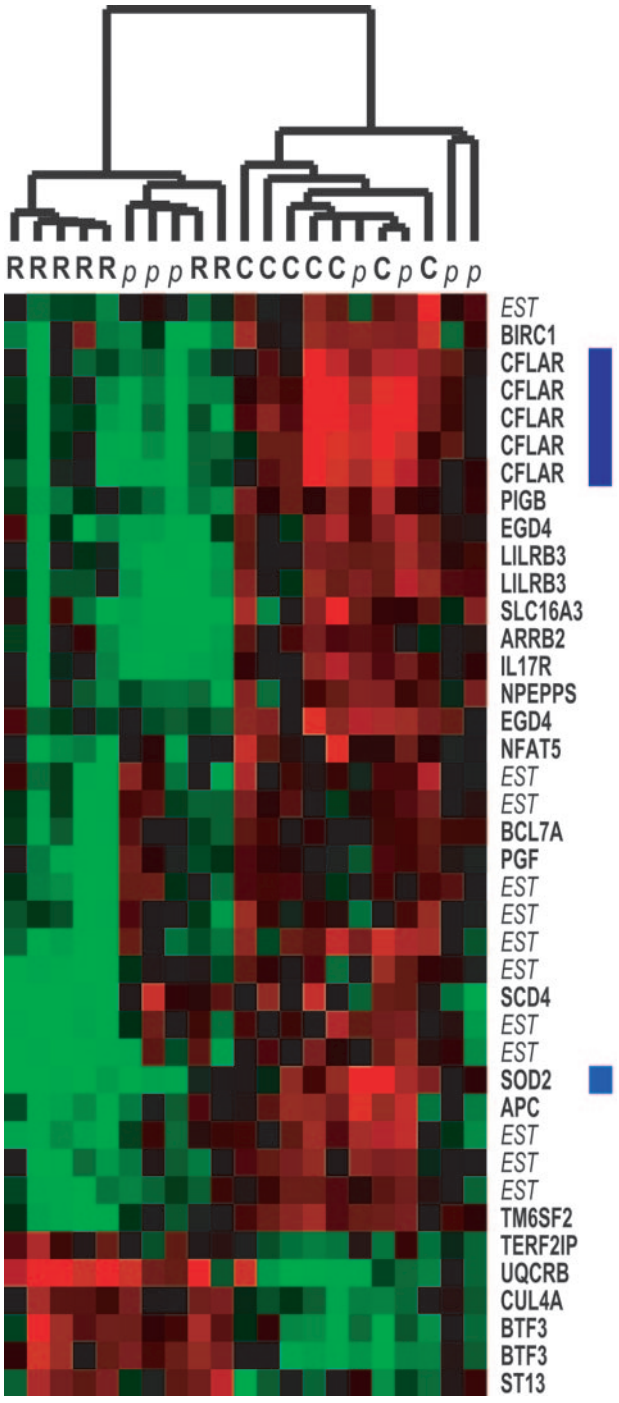


Figure 3. Cluster analysis. We analyzed 40 candidate markers with hierarchical clustering (see Table 2 for full names and functional annotation of 40 candidates). Results are displayed with an Eisen plot, which consists of a dendrogram to demonstrate relationships among samples and color-coded heat map to display level of expression of individual genes. For each gene, red indicates higher-than-median expression, and green indicates lower-than-median expression. As shown in the dendrogram, our candidate markers partition rejection (R) and control (C) samples into 2 main branches. Postrejection samples (*p*) are present in both main branches, which indicates intermediate expression profiles for this group. Genes chosen for subsequent qRT-PCR validation are indicated with blue squares.

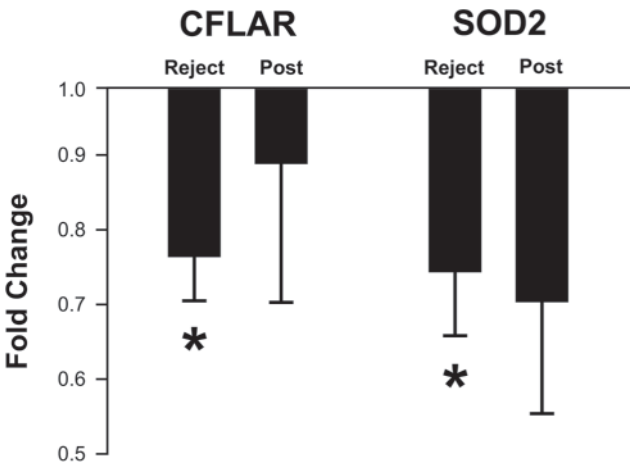


Figure 4. We quantified transcript abundance of 2 candidate markers, *CFLAR* and *SOD2*, using qRT-PCR. Data are displayed as fold changes in expression in rejection (n=10) and postrejection (n=8), each compared with control (n=5). In agreement with microarray findings, both *CFLAR* and *SOD2* expression were decreased in rejection. *CFLAR* expression returned toward control levels in postrejection samples, and *SOD2* expression remained low, consistent with persistent partial activation of circulating leukocytes after treatment of rejection. *$P<0.05$ compared with control by Wilcoxon rank-sum test.

ple genes, which integrates small changes in a large number of component markers, will prove to be the most robust diagnostic approach for detecting allograft rejection noninvasively.

In addition to viewing our candidate genes as diagnostic markers of rejection, a portion of them may mediate components of rejection. As shown in Table 2, the known or proposed function of our candidates involves cellular processes that are plausible components of an immune response, such as regulation of DNA transcription or translation, cell-cycle and apoptosis regulators, and markers of immune system activation. It is possible that changes in expression of genes involved in the regulation of programmed cell death, such as *CFLAR*, promote clonal expansion of specific lymphocyte populations as part of the rejection process[26]; however, our ability to make specific biological inferences is limited by the mixed cell population examined and the observational nature of the present study. Determining which, if any, of our candidate genes contribute to rejection will require experimental approaches.

Expression profiling is a powerful technique, but it creates substantial challenges that result from the analysis of many genes in a small number of samples. We addressed these concerns at multiple levels. First, we used conservative normalization and gene-selection strategies that are superior in the analysis of relatively small data sets.[15] Second, we incorporated serial measurements in the same patients, which reduces the impact of interpatient variability.[27] Third, we validated selected findings using quantitative PCR. Fourth, we performed our analyses on immediately preserved whole-blood isolates, which minimizes the impact of sample preprocessing procedures, such as cell sorting or buffy coat isolation, on the gene-expression profile and is more convenient to implement in a clinical setting.[28] The major limitation

of this proof-of-principle study is the small sample size, which limits our ability to assess the influence of confounding factors, such as age, on changes in peripheral gene expression.

In conclusion, we have demonstrated the principle that peripheral blood gene expression correlates with cardiac allograft rejection. Further studies are necessary to test our panel of markers prospectively with the goal of developing a clinically useful, noninvasive test for cardiac allograft rejection.

Acknowledgments

This work was supported by National Institutes of Health grant K23HL071562 and an allocation from the Commonwealth of Pennsylvania.

References

1. Billingham ME, Cary NR, Hammond ME, Kemnitz J, Marboe C, McCallister HA, Snovar DC, Winters GL, Zerbe A. A working formulation for the standardization of nomenclature in the diagnosis of heart and lung rejection: Heart Rejection Study Group: the International Society for Heart Transplantation. *J Heart Transplant*. 1990;9:587–593.
2. Hosenpud JD, Bennett LE, Keck BM, Boucek MM, Novick RJ. The Registry of the International Society for Heart and Lung Transplantation: seventeenth official report: 2000. *J Heart Lung Transplant*. 2000;19:909–931.
3. Pophal SG, Sigfusson G, Booth KL, Bacanu SA, Webber SA, Ettedgui JA, Neches WH, Park SC. Complications of endomyocardial biopsy in children. *J Am Coll Cardiol*. 1999;34:2105–2110.
4. de Groot-Kruseman HA, Baan CC, Hagman EM, Mol WM, Niesters HG, Maat AP, Zondervan PE, Weimar W, Balk AH. Intragraft interleukin 2 mRNA expression during acute cellular rejection and left ventricular total wall thickness after heart transplantation. *Heart*. 2002;87:363–367.
5. Shulzhenko N, Morgun A, Rampim GF, Franco M, Almeida DR, Diniz RV, Carvalho AC, Gerbase-DeLima M. Monitoring of intragraft and peripheral blood TIRC7 expression as a diagnostic tool for acute cardiac rejection in humans. *Hum Immunol*. 2001;62:342–347.
6. Shulzhenko N, Morgun A, Zheng XX, Diniz RV, Almeida DR, Ma N, Strom TB, Gerbase-DeLima M. Intragraft activation of genes encoding cytotoxic T lymphocyte effector molecules precedes the histological evidence of rejection in human cardiac transplantation. *Transplantation*. 2001;72:1705–1708.
7. Shulzhenko N, Morgun A, Franco M, Souza MM, Almeida DR, Diniz RV, Carvalho AC, Pacheco-Silva A, Gerbase-Delima M. Expression of CD40 ligand, interferon-gamma and Fas ligand genes in endomyocardial biopsies of human cardiac allografts: correlation with acute rejection. *Braz J Med Biol Res*. 2001;34:779–784.
8. van Emmerik N, Baan C, Vaessen L, Jutte N, Quint W, Balk A, Bos E, Niesters H, Weimar W. Cytokine gene expression profiles in human endomyocardial biopsy (EMB) derived lymphocyte cultures and in EMB tissue. *Transpl Int*. 1994;7(suppl 1):S623–S626.
9. Alpert S, Lewis NP, Ross H, Fowler M, Valantine HA. The relationship of granzyme A and perforin expression to cardiac allograft rejection and dysfunction. *Transplantation*. 1995;60:1478–1485.
10. Baan CC, van Emmerik NE, Balk AH, Quint WG, Mochtar B, Jutte NH, Niesters HG, Weimar W. Cytokine mRNA expression in endomyocardial biopsies during acute rejection from human heart transplants. *Clin Exp Immunol*. 1994;97:293–298.
11. Kobashigawa JA. Treatment of nonhemodynamic compromising rejection: conventional approaches vs individualization/new immunosuppressive drugs. *Transplant Proc*. 1997;29:37S–39S.
12. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A*. 2001;98:31–36.
13. Seo J, Bakay M, Chen YW, Hilmer S, Shneiderman B, Hoffman EP. Interactively optimizing signal-to-noise ratios in expression profiling: project-specific algorithm selection and detection p-value weighting in Affymetrix microarrays. *Bioinformatics*. 2004;20:2534–2544.
14. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. 2003;4:249–264.
15. Wu Z, Irizarry RA. Preprocessing of oligonucleotide array data. *Nat Biotechnol*. 2004;22:656–658.
16. Ihaka R, Gentleman R. A language for data analysis and graphics. *J Graphic Comput Stat*. 1996;5:299–314.
17. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A*. 2001;98:5116–5121.
18. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A*. 1998;95:14863–14868.
19. Kimball PM, Radovancevic B, Isom T, Spickard A, Frazier OH. The paradox of cytokine monitoring-predictor of immunologic activity as well as immunologic silence following cardiac transplantation. *Transplantation*. 1996;61:909–915.
20. Lagoo AS, George JF, Naftel DC, Griffin AK, Kirklin JK, Lagoo-Deenadayalan S, Hardy KJ, Savunen T, McGiffin DC. Semiquantitative measurement of cytokine messenger RNA in endomyocardium and peripheral blood mononuclear cells from human heart transplant recipients. *J Heart Lung Transplant*. 1996;15:206–217.
21. Morgun A, Shulzhenko N, Diniz RV, Almeida DR, Carvalho AC, Gerbase-DeLima M. Cytokine and TIRC7 mRNA expression during acute rejection in cardiac allograft recipients. *Transplant Proc*. 2001;33:1610–1611.
22. Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, Beverloo HB, Moorhouse MJ, van der Spek PJ, Lowenberg B, Delwel R. Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med*. 2004;350:1617–1628.
23. Bullinger L, Dohner K, Bair E, Frohling S, Schlenk RF, Tibshirani R, Dohner H, Pollack JR. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med*. 2004;350:1605–1616.
24. Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, Gascoyne RD, Muller-Hermelink HK, Smeland EB, Giltnane JM, Hurt EM, Zhao H, Averett L, Yang L, Wilson WH, Jaffe ES, Simon R, Klausner RD, Powell J, Duffey PL, Longo DL, Greiner TC, Weisenburger DD, Sanger WG, Dave BJ, Lynch JC, Vose J, Armitage JO, Montserrat E, Lopez-Guillermo A, Grogan TM, Miller TP, LeBlanc M, Ott G, Kvaloy S, Delabie J, Holte H, Krajci P, Stokke T, Staudt LM. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med*. 2002;346:1937–1947.
25. Sarwal M, Chua MS, Kambham N, Hsieh SC, Satterwhite T, Masek M, Salvatierra O Jr. Molecular heterogeneity in acute renal allograft rejection identified by DNA microarray profiling. *N Engl J Med*. 2003;349:125–138.
26. Inohara N, Koseki T, Hu Y, Chen S, Nunez G. CLARP, a death effector domain-containing protein interacts with caspase-8 and regulates apoptosis. *Proc Natl Acad Sci U S A*. 1997;94:10717–10722.
27. The Tumor Analysis Best Practices Working Group. Expression profiling: best practices for data generation and interpretation in clinical trials. *Nat Rev Genet*. 2004;5:229–237.
28. Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Moser K, Ortmann WA, Espe KJ, Balasubramanian S, Hughes KM, Chan JP, Begovich A, Chang SY, Gregersen PK, Behrens TW. Expression levels for many genes in human peripheral blood cells are highly sensitive to ex vivo incubation. *Genes Immun*. 2004;5:347–353.

American Journal of Transplantation 2006; 6: 150–160
Blackwell Munksgaard

doi: 10.1111/j.1600-6143.2005.01175.x

Noninvasive Discrimination of Rejection in Cardiac Allograft Recipients Using Gene Expression Profiling

M. C. Deng*, H. J. Eisen*, M. R. Mehra, M. Billingham, C. C. Marboe, G. Berry, J. Kobashigawa, F. L. Johnson, R. C. Starling, S. Murali, D. F. Pauly, H. Baron, J. G. Wohlgemuth, R. N. Woodward, T. M. Klingler, D. Walther, P. G. Lal, S. Rosenberg and S. Hunt for the CARGO Investigators

Drexel University College of Medicine, Philadelphia, Pennsylvania, USA (HJE), Columbia University, New York, NY, USA (MCD, CM, HB), University of Maryland, Baltimore, Maryland, USA (MM, FLJ), Stanford University School of Medicine, Stanford, California, USA (MB, GB, SH), University of California at Los Angeles, Los Angeles, California, USA (JK), Cleveland Clinic Foundation, Cleveland, Ohio, USA (RCS), University of Pittsburgh, Pennsylvania, USA (SM), University of Florida, Gainesville, Florida, USA (DP), XDx, Inc., South San Francisco, California, USA (JGW, RNW, TMK, SR, DW, PGL)
M. C. Deng and H. J. Eisen contributed equally to this work. Corresponding authors: H. J. Eisen, Howard.Eisen@DrexelMed.edu; Mario C. Deng, md785@columbia.edu*

Rejection diagnosis by endomyocardial biopsy (EMB) is invasive, expensive and variable. We investigated gene expression profiling of peripheral blood mononuclear cells (PBMC) to discriminate ISHLT grade 0 rejection (quiescence) from moderate/severe rejection (ISHLT $\geq$3A). Patients were followed prospectively with blood sampling at post-transplant visits. Biopsies were graded by ISHLT criteria locally and by three independent pathologists blinded to clinical data. Known alloimmune pathways and leukocyte microarrays identified 252 candidate genes for which real-time PCR assays were developed. An 11 gene real-time PCR test was derived from a training set (n = 145 samples, 107 patients) using linear discriminant analysis (LDA), converted into a score (0–40), and validated prospectively in an independent set (n = 63 samples, 63 patients). The test distinguished biopsy-defined moderate/severe rejection from quiescence (p = 0.0018) in the validation set, and had agreement of 84% (95% CI 66% C94%) with grade ISHLT $\geq$3A rejection. Patients $>$1 year post-transplant with scores below 30 (approximately 68% of the study population) are very unlikely to have grade $\geq$3A rejection (NPV = 99.6%). Gene expression testing can detect absence of moderate/severe rejection, thus avoiding biopsy in certain clinical settings. Additional clinical experience is needed to establish the role of molecular testing for clinical event prediction and immunosuppression management.

Key words: Allogeneic, biological markers, gene expression profiling, graft rejection, heart transplantation, immune response genes, immunologic, immunologic monitoring, transplantation.

Received 22 April 2005, revised and accepted for publication 7 October 2005

Introduction

The goal of care after cardiac transplantation is to prevent allograft rejection while minimizing immunosuppressive side effects (1,2). The gold standard of rejection surveillance in cardiac transplantation is endomyocardial biopsy (EMB). However, EMB is invasive, expensive, subject to sampling error, inter-observer variability (3–5), and causes morbidity (0.5–1.5%). Although noninvasive alternatives to EMB are clearly needed, methods such as echocardiography, ultrasonic myocardial back-scatter, radionuclide imaging, magnetic resonance imaging, intra-myocardial electrograms and multiparametric immune monitoring have been difficult to validate and implement (6–20).

As recirculating peripheral blood mononuclear cells (PBMC) may reflect earlier host responses than those at local sites, measurement of PBMC gene expression might provide useful diagnostic information and reduce the need for EMB in patients who are asymptomatic. Recent studies using microarray analysis (21) or real-time PCR analysis of cytokine genes (22) have suggested that gene expression measurements in PBMC may be correlated with cardiac allograft rejection. However, these single center studies are limited by the absence of methodology to recognize the imperfect 'gold standard' nature of EMB, which creates significant challenges for diagnostic development and validation study design and analysis (23,24). In addition, the absence of multicenter independent validation sets in both studies suggests the need for more extensive investigation.

Based on the assumption that a gene expression signature of immune activation and leukocyte trafficking would be detectable in recipient PBMC and reflect the rejection status of the donor allograft, we tested the hypothesis that

a gene expression test could discriminate ISHLT grade 0 rejection (quiescence) from moderate/severe (ISHLT grade ≥3A) rejection (nonquiescence).

Methods

Study design

After approval by local Institutional Reviews Boards at eight centers, all patients undergoing heart transplantation and providing informed consent were eligible for the Cardiac Allograft Rejection Gene Expression Observational (CARGO) study beginning in September 2001. Enrolled patients were followed at each clinical encounter with data collection including EMB, hemodynamics and/or echocardiography, immunosuppression, laboratory data and complications, which were captured in electronic clinical report forms. EMB slides were obtained from centers for interpretation by a panel of pathologists blinded to the clinical data.

The study was conducted in three phases (Figure 1A): (1) *candidate gene discovery* using a combination of focused genomic and knowledge-base approaches; (2) *diagnostic development* using PCR assays and rigorous statistical methods and (3) *validation* in a prospective and blinded study. Samples were selected and divided into a training set, used for candidate gene discovery and diagnostic development, and a set used for validation of the gene expression signature described below.

Data from an additional set of representative samples not used in any of the three phases were evaluated after the validation studies to estimate the negative predictive value (NPV) and positive predictive value (PPV) in the CARGO population.

Biopsy Samples

Biopsies performed by standard techniques were graded by local pathologists. A subset of biopsy samples, including all local grades 1B, 2, 3A and 3B and a representative set of grades 0 and 1A samples, were also graded by three independent ('central') pathologists blinded to clinical information. After an evaluation of the concordance of these biopsy grades by the four pathologists, criteria for selecting acute cellular *rejection* and *quiescent* samples were defined prior to developing and validating the classifier.

Blood Samples

PBMC were isolated from eight mL of venous blood using density gradient centrifugation (CPT, Becton-Dickinson). Samples were frozen in lysis buffer (RLT, Qiagen) within 2 h of phlebotomy. Total RNA was isolated from each sample (RNeasy, Qiagen).

The effects of processing time on gene assays were tested in PBMC isolated from six venous blood samples from each of nine donors. Samples were treated identically except the interval between blood draw and first centrifugation step was varied from 1 to 8 h. Any gene assays showing significant systematic variations across this time period were eliminated from the development process.

Candidate Gene Discovery

Microarray expression profiling

A custom microarray was designed using RNA sequences expressed in stimulated and resting human leukocytes (PCR Select, Clontech) and from publicly available sequence databases. A total of 7370 genes were represented by 50-mer oligonucleotides (Sigma) on a spotted custom microarray

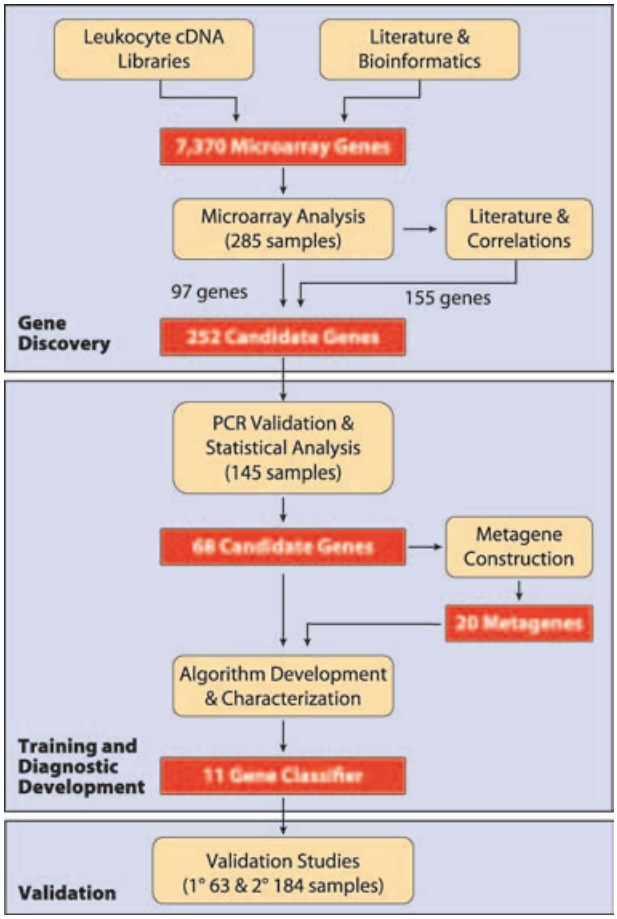


Figure 1: (A) Strategy used for the gene discovery, diagnostic development and validation. Initial discovery efforts using genomic (microarray) and knowledge-base (literature and sequence analyses) method produced 252 candidate genes for further quantitative PCR assay development. These 252 assays were applied to 145 samples to generate the dataset for training of the classifier. Statistical learning methods were used with these real-time PCR measurements to refine the gene set to 68 genes. Gene expression correlations were captured as composite metagenes. Automated statistical methods were used to build a classifier, and bootstrap and cross-validation methods to estimate classifier performance. The linear discriminant classifier developed was evaluated on an independent, blinded validation set from independent patients (primary validation set) and independent samples (secondary validation set).

(Telechem). To increase the power and quality of results, a large number of clinical samples (285) were used. Microarray data are available at GEO (25) with accession number GSE2445. The experimental methods are described in detail in the Supplement Section.

Knowledge-base gene discovery

Our leukocyte-focused genomic microarray approach was complemented with (1) a review of the literature on pathways involved in immune activation, recruitment and mobilization, in general, and solid organ transplant rejection, in particular; and (2) genes related to genes suggested to be significant by microarrays (by pathways and families).

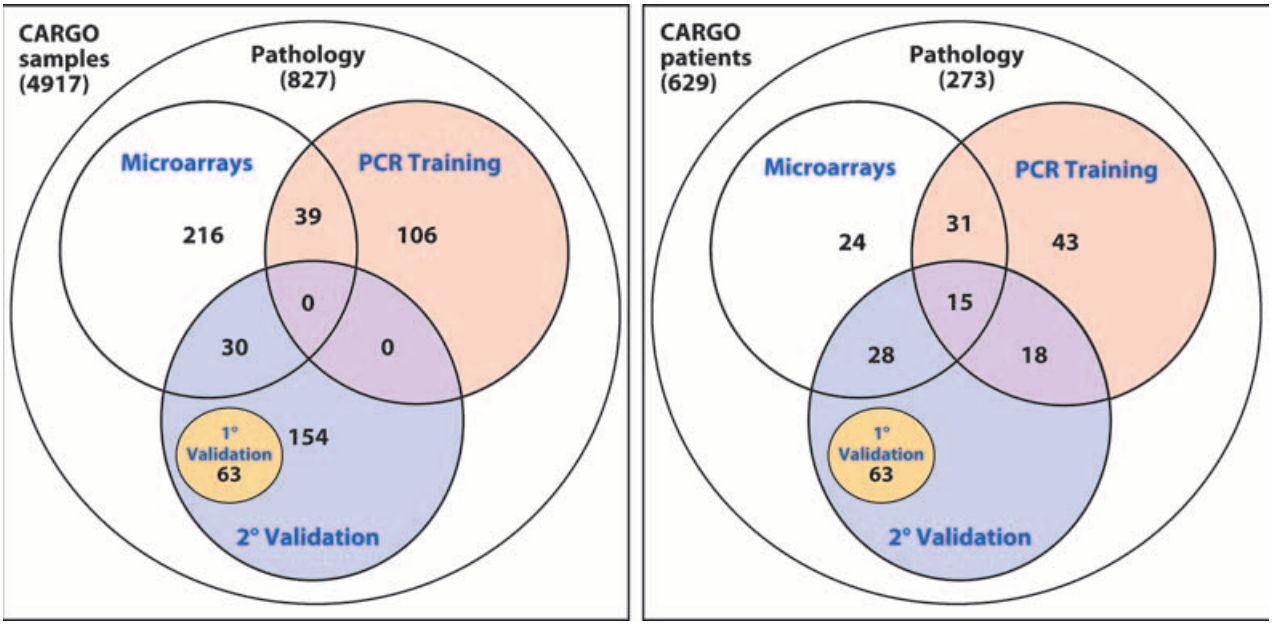


Figure 1: (B) Venn diagrams of samples and patients used for CARGO studies. A total of 827 samples were examined by centralized pathology. A set of samples and patients independent from both the microarray and PCR training studies are reported in the primary validation study, whereas an independent set of samples from the PCR training was used in the secondary validation study.

Diagnostic Development

Real-time PCR (RT-PCR)

PCR primers and probes were designed using PRIMER3 (version 0.9, Whitehead Research Institute). Assays were designed on the full-length mRNA after masking to avoid problematic sequences. Assays were qualified for inclusion in the training set by specificity, linear dynamic range and efficiency using both human PBMC cDNA and synthetic oligonucleotide templates. For each gene, triplicate 10 μL real-time PCR reactions were performed on the ABI 7900HT system using FAM-TAMRA probes and standard Taqman protocols (Applied Biosystems) on cDNA from 0.5 ng total RNA.

Normalization and control genes

Normalization genes were empirically selected using PCR data from the training samples. Genes which did not discriminate between rejection and quiescence samples with small standard deviations across all samples were considered as normalization genes. Six such genes spanning different expression levels were chosen.

Three additional assays were included as controls: two to detect genomic DNA contamination by the difference between a transcribed and nontranscribed region of the Gus-B gene and the third, a spiked-in control template for an Arabidopsis gene to determine if the PCR reaction was successful.

Discriminant equation development

RT-PCR data on 252 genes for the training set of 36 rejection and 109 quiescent CARGO samples were generated to derive a panel of candidate genes for classifier development and to validate microarray results. Gene expression results were analyzed with Student's *t*-test, median ratios, hierarchical clustering by TreeView and an expert assessment of biological relevance. Metagenes, defined as transcripts behaving in a concordant manner (26), were constructed by averaging gene expression levels that were correlated across training samples with correlation coefficients of at least 0.7. Genes significantly distinguishing rejection from quiescence in the PCR training set by *t*-test ($p \leq 0.01$), by median ratio differences of <0.75 or >1.25 or by correlation with significant genes were used for metagene construction and classifier development.

The methods for analyzing gene expression data included principal components analysis, linear discriminant analysis (LDA, StatSoft, Inc.), logistic regression (SAS Institute, Inc.), prediction analysis of microarrays (PAM) (27), voting, classification and regression trees (TreeNet, Salford Systems), Random Forests, nearest shrunken centroids and k-nearest neighbors. We sought to develop a classifier that quantitatively distinguished current moderate/severe acute cellular rejection (ISHLT grade ≥3A) from quiescence (ISHLT grade 0) using gene and metagene expression levels as the variables.

The final classifier was developed using LDA as implemented in the 'discriminant function analysis' module of Statistica (StatSoft, Inc.). LDA constructs a linear classifier by automatically selecting genes and/or metagenes that, in combination, optimally separate rejection and quiescent samples in the training set. The robustness of selected genes and the appropriate number of genes in the classifier were both evaluated by cross-validation.

Validation

Design

An independent cohort of CARGO patients was selected to validate the effectiveness of the LDA classifier defined in the diagnostic development phase using a prospective and blinded study protocol. The primary objective of the validation study was to test the pre-specified hypothesis that the diagnostic score distinguishes quiescence, defined as ISHLT grade 0, from moderate/severe biopsy-proven acute rejection, defined as ISHLT grade ≥3A, both grades determined from local and centralized cardiopathological examination. This was assessed using a 2-tailed Student's

t-test for comparing score distributions for rejection and quiescent samples. Secondary and exploratory objectives included documentation of diagnostic performance across thresholds and description of correlations to clinical variables.

Results for the validation study are reported for unique samples from patients not used for training (primary validation study), as well as for a larger set of samples not used for training (secondary validation study). These latter samples may provide improved power but may be biased to the extent that a longitudinal set of samples from an individual patient are not completely independent with respect to gene expression.

Prevalent Population Studies

A representative set of samples, across all local biopsy grades and ≥1 year post-transplant were evaluated to assess the discriminant equation performance on a stable patient population. From these samples, PPV (fraction of samples with scores at or above the threshold expected to have concurrent biopsy grade ≥3A) and NPV (fraction of samples with scores below the threshold expected to be free from biopsy grade ≥3A) were estimated at multiple test thresholds. Given the risk associated with undetected acute cellular rejection, and the clinical use of EMB, we sought a threshold that maximized the NPV at the expense of the PPV.

Quantitative CMV assays

Plasma was tested for quantitative CMV viral load using the COBAS protocol (Roche). These samples were selected from the CARGO study and represented known or suspected CMV infection and matched controls.

Results

Patients and samples used in these studies were selected from the CARGO database, with donor and recipient characteristics similar to those reported by the United Network for Organ Sharing (UNOS) for 2003 (28) (Table 1). The relationships between the samples and patients used in the three phases are shown in Figure 1B.

Gene discovery

In the gene discovery phase, 285 rejection and quiescent samples from 98 patients were hybridized to the leukocyte microarrays covering 7370 genes. Ninety-seven genes were selected as candidates for PCR assay development from these microarray studies based on false detection rates from SAM <20% (29), p-values in nonparametric analysis <0.05 or clustering with genes involved in rejection. This gene set was expanded to include related genes identified by correlated expression or related functions, as well as genes from the literature involved in transplant rejection, yielding an additional 155 gene candidates.

Diagnostic development

In the diagnostic development phase, 252 real-time PCR assays were developed to assess and confirm the discriminatory ability of the candidate genes from the gene discovery phase. These PCR assays were performed on 145 samples including 36 rejections (from 28 patients) and 109 quiescent samples (from 86 patients). Centralized pathology reading was used to identify these samples, where at least two of four pathologists were required to classify a sample as grade ≥3A for rejection, and three of four pathologists were required to classify a sample as grade 0 for quiescence. These criteria were set prospectively based upon centralized reading of over 800 CARGO samples and were used in the diagnostic development and validation PCR studies (30).

Analysis of this set of PCR data (see PCR-heatmap Figure 2A) yielded 68 genes that distinguished rejection samples from quiescent samples by *t*-test ($p < 0.01$), median ratio of (>1.25 or <0.75), or by correlation to discriminatory genes (Table 2). By hierarchical clustering (31) (Figure 2B), the predominant genes showing increased expression with rejection were T-cell/NK and $CD8^+$ T-cell activation markers (perforin, granulysin) and erythropoiesis markers (ALAS2, WDR40A, MIR). Six genes (CXCR4, hIAN7, HBG, CXCR3, ADM and TNFSF6) were eliminated due to significant variation in gene expression with sample processing time (32) yielding 62 genes for discriminatory signature development.

To take advantage of the correlations observed in gene expression (Figure 2B), 20 metagenes (26) were created by averaging correlated gene expression levels of the 62 genes. Using the training data set of 145 samples and these 82 variables (62 genes and 20 metagenes), a linear discriminant equation was derived by sequentially fitting the gene expression data to maximize agreement with the biopsy-based samples classification. The final equation, yielding a score between 0 and 40, combines the expression levels of four individual genes and three metagenes, constituting 11 genes in total (five from microarray and six from literature), which best distinguished rejection from quiescence (Table 2, Figure 2B). Additional terms did not further improve performance above approximately 75% correct classification. More complex statistical methods than LDA did not yield better performance and are less amenable to rigorous cross-validation as seen in other analyses (33).

Validation

In the validation phase, the discriminant equation performance was first estimated using the bootstrap method on the entire training set of samples. As shown in Table 3, agreement with biopsy ≥3A and biopsy grade 0, was estimated at 80% and 59%, respectively, at a single, predefined threshold of 20 (scores ≥20 indicate rejection).

To rigorously determine the test performance, an independent primary validation set of 63 unique samples (31 rejection, 32 quiescent samples) from 63 patients was tested in a prospective and blinded manner. The classifier distinguished moderate/severe rejection from quiescence (*t*-test, $p = 0.0018$). At the prospectively defined score threshold of 20, 84% (95% CI 66–94%) of

Table 1: Clinical characteristics of study patients and samples

			Microarray discovery			Diagnostic development & training			PCR 1° validation (unique patients)			PCR 2° validation (unique samples)		
	UNOS 2003	CARGO (N = 629 patients, 4917 samples)	Rejection (N = 28 patients, 38 samples)	No rejection (N = 94 patients, 247 samples)	p-value	Rejection (N = 28 patients, 36 samples)	No rejection (N = 86 patients, 109 samples)	p-value	Rejection (N = 31 patients, 31 samples)	No rejection (N = 32 patients, 32 samples)	p-value	Rejection (N = 50 patients, 62 samples)	No rejection (N = 83 patients, 122 samples)	p-value
Recipient age														
Under 18	14.0%	6.4%	0.0%	0.0%		0.0%	1.8%		6.4%	12.5%		3.2%	1.6%	
18–34	9.4%	9.9%	15.8%	13.0%		13.9%	5.5%		12.9%	9.4%		14.5%	13.9%	
35–49	21.1%	17.7%	23.7%	12.6%	NS	13.9%	21.1%	NS	3.2%	18.7%	0.006	17.7%	18.0%	NS
50–64	47.2%	53.1%	57.9%	65.6%		69.4%	56.9%		67.7%	31.2%		53.2%	54.9%	
65+	8.5%	12.9%	2.6%	8.9%		2.8%	14.7%		9.7%	28.1%		11.3%	11.5%	
Recipient race														
White	71.1%	72.3%	73.7%	74.1%		72.2%	78.9%		70.9%	65.6%		67.7%	66.4%	
Black	16.0%	17.3%	21.1%	15.4%		19.4%	10.1%		16.1%	15.6%		24.2%	18.9%	
Hispanic	8.8%	6.2%	5.3%	7.7%	NS	8.3%	6.4%	NS	6.4%	3.1%	NS	6.5%	9.0%	NS
Asian	1.9%	1.2%	0.0%	0.0%		0.0%	1.8%		3.2%	3.1%		0.0%	2.5%	
Other	2.1%	3.1%	0.0%	2.8%		0.0%	2.8%		0.0%	3.1%		1.6%	3.3%	
Recipient sex														
Male	73.6%	74.6%	92.1%	77.3%	NS	86.1%	73.4%	NS	74.2%	78.1%	NS	80.6%	81.1%	NS
Female	26.4%	25.4%	7.9%	22.7%		13.9%	26.6%		22.6%	12.5%		19.4%	18.9%	
Donor age														
Under 18	21.6%	17.9%	18.4%	11.4%		17.1%	17.5%		16.1%	21.8%		20.0%	17.1%	
18–34	44.2%	46.5%	47.4%	49.4%		51.4%	46.6%		45.2%	46.8%		38.2%	51.3%	
35–49	25.7%	24.9%	26.3%	28.2%	NS	22.9%	24.3%	NS	6.45%	15.6%	NS	32.7%	19.7%	NS
50–64	8.2%	10.5%	7.9%	11.0%		8.6%	11.7%		16.9%	9.3%		9.1%	11.1%	
65+	0.2%	0.2%	0.0%	0.0%		0.0%	0.0%		0.0%	3.1%		0.0%	0.9%	
Donor race														
White	69.6%	71.4%	73.7%	66.9%		75.0%	62.4%		86.6%	67.8%		83.9%	66.4%	
Black	12.0%	14.2%	5.3%	14.6%		5.6%	12.8%		10%	14.3%		8.1%	13.9%	
Hispanic	15.7%	11.4%	21.1%	16.3%	NS	19.4%	18.3%	NS	3.3%	14.3%	NS	8.1%	13.1%	NS
Asian	1.6%	0.7%	0.0%	2.1%		0.0%	0.9%		0.0%	0.0%		0.0%	0.8%	
Other	1.2%	2.3%	0.0%	0.0%		0.0%	5.5%		0.0%	3.5%		0.0%	5.7%	
Donor sex														
Male	68.4%	63.7%	71.1%	66.1%	NS	61.1%	68.5%	NS	53.3%	75%	NS	58.1%	74.2%	NS
Female	31.6%	36.3%	28.9%	33.9%		38.9%	31.5%		46.7%	25%		41.9%	25.8%	
Primary diagnosis														
Coronary artery disease	42.1%	23.8%	15.8%	21.9%		22.2%	32.1%		29.1%	34.4%		16.1%	30.3%	
Cardiomyopathy	47.0%	70.0%	81.6%	69.2%		77.8%	58.7%		61.3%	53.1%		79.0%	64.8%	
Congenital heart disease	8.5%	2.3%	0.0%	0.4%	NS	0.0%	1.8%	NS	0.0%	3.1%	NS	0.0%	2.5%	0.025
Retransplant	3.3%	0.5%	2.6%	5.3%		0.0%	1.8%		0.0%	0.0%		3.2%	0.8%	
Valvular disease	1.9%	2.4%	0.0%	1.2%		0.0%	1.8%		0.0%	0.0%		0.0%	0.8%	
Other	0.5%	1.0%	0.0%	2.0%		0.0%	3.7%		6.4%	0.0%		1.6%	0.8%	
Immunosuppression*														
Cyclosporine	64.9%	50.4%	71.1%	71.7%		52.8%	44.0%		58.1%	34.4%		72.6%	53.3%	
FK-506	42.9%	36.6%	26.3%	25.5%		47.2%	54.1%		38.7%	50%		25.8%	38.5%	
Mycophenolate	80.5%	72.0%	81.6%	87.4%		72.2%	78.0%		74.2%	75%		80.6%	83.6%	
Rapamycin	7.5%	9.8%	5.3%	2.4%	NS	22.2%	14.7%	NS	12.9%	6.3%	0.048	12.9%	8.2%	NS
Azathioprine	14.7%	1.4%	0.0%	0.4%		0.0%	1.8%		0.0%	3.1%		0.0%	1.6%	
Corticosteroids	91.1%	82.2%	97.4%	94.3%		94.4%	91.7%		77.4%	75%		88.7%	82.8%	
Zenapax	2.5%	9.4%	2.6%	4.5%		0.0%	7.3%		3.2%	3.1%		6.5%	4.1%	
Days post-Tx														
Average days post-Tx	NA	241	83	62		254	206		149	147		205	265	

NS = Not significant ($p \geq 0.05$).

*From UNOS 2001 data. This percentage represents the number of transplants in which a particular drug was used for maintenance at any point in the year after transplant divided by the number of transplants in 2001, and only accounts for patients with immunosuppressive information.

Comparison of clinical parameters of patients and samples used in the microarray, diagnostic development and validation studies.

CARGO = Cardiac Allograft Rejection Gene expression Observation study.

UNOS = United Network for Organ Sharing.

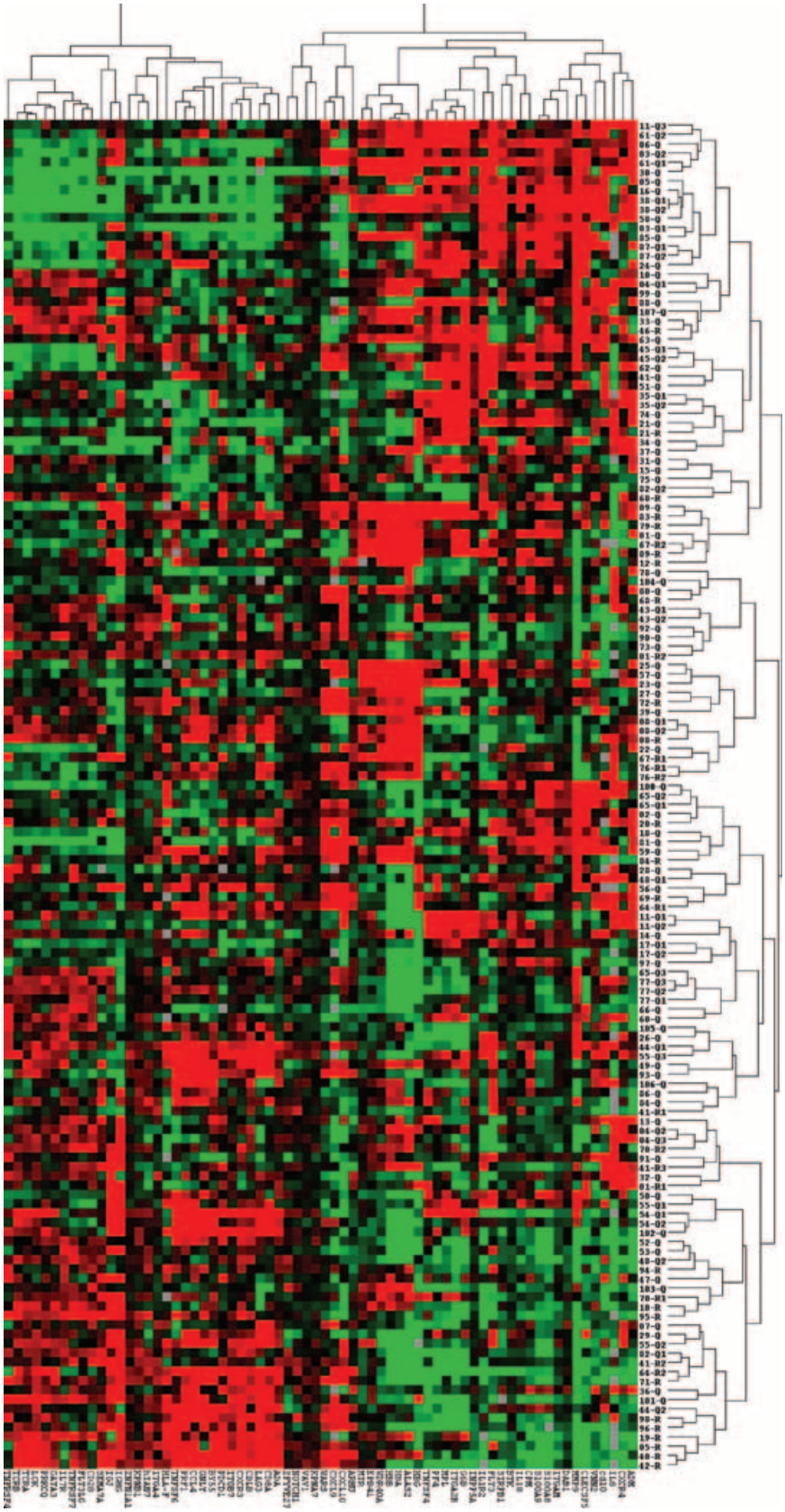


Figure 2: (A) Heat map of PCR expression measurements for 68 genes listed in Table 2 for 145 training samples. A heat map (red = up-regulated; green = down-regulated) for genes that significantly (*t*-test $p < 0.01$) differ between rejection and quiescence or have median ratios >1.25 or <0.75 or are correlated with significant genes is shown. Complete linkage clustering was used with Pearson correlation as the distance metric: genes (labeled by gene IDs) are clustered vertically according to correlated expression across samples; samples (labeled by patient no.— rejection (R) or quiescent (Q) and an ordinal if multiple samples from a single patients) are clustered horizontally according to correlated expression across genes.

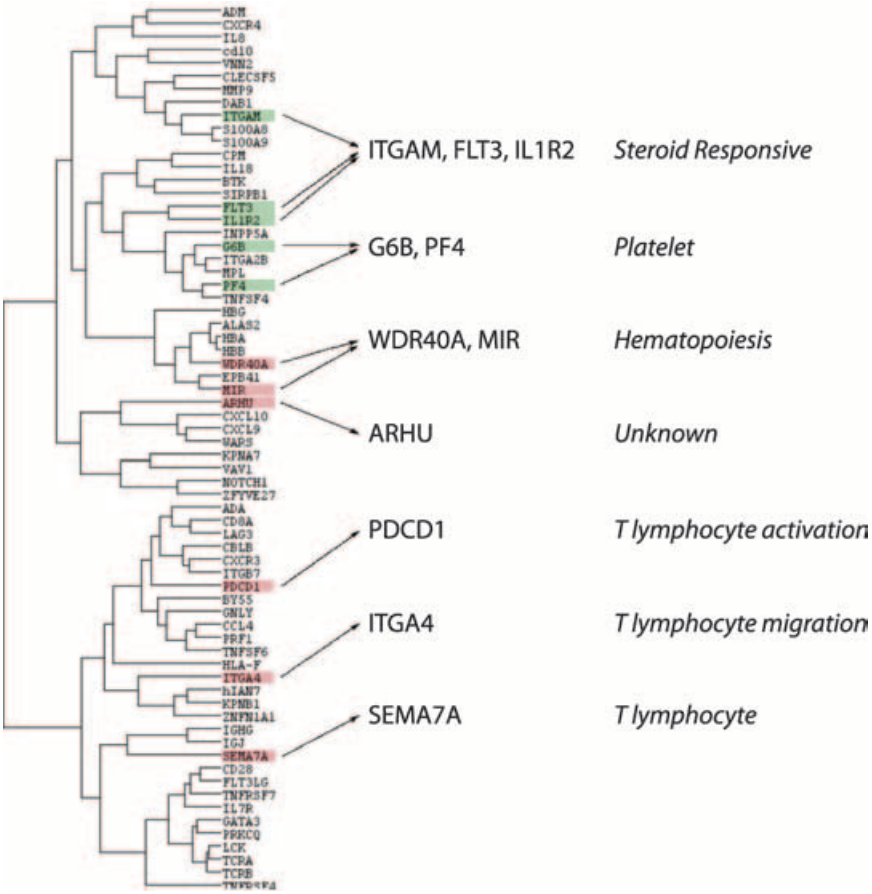


Figure 2: (B) Dendrogram of genes listed in Table 2, showing algorithm composition and cell type/function annotation. Gene correlation tree from heat map in Figure 2B expanded to show up- and down-regulated genes included in discriminant equation. The equation consists of a constant, 3 metagenes (of 3, 2 and 2 genes, respectively) and 4 single gene terms. Terms are annotated with functional, pathway or cell-type information where the supporting biology is known.

rejection and 38% (95% CI 22–56%) of quiescence samples were classified correctly. Receiver operating characteristics (ROC) analysis, shown in Figure 3A, yields an area under the curve of 0.72 ± 0.06. Similar results were obtained for the secondary validation set of 184 samples (62 rejection, 122 quiescent) from 124 patients that includes the 63 primary validation patient samples and additional samples from patients who contributed samples to the gene discovery or diagnostic development phases of this work (Figure 1B). With the same threshold of 20, 76% (95% CI 63–85% of rejections and 41% (95% CI 32–50%) of quiescent samples were classified correctly ($p = 0.0001$).

Analysis of the validation studies showed that time post-transplant was the single most important score-correlated variable. This time-dependence was responsible for the overall low specificity relative to biopsy using the single score threshold of 20. Scores increased with time post-transplant in association with the weaning of maintenance steroid doses, which generally occurs in the first year. Therefore, we investigated performance relative to biopsy in the ≥6 months and ≥1 year post-transplant periods (Table 3). In the ≥6 months period with a threshold of 28, 71% of rejection and 79% of quiescent secondary validation samples were classified correctly. For ≥1 year post-transplant, a threshold of 30 results in 80% of rejection and 78% of quiescent samples classified correctly. Similar improvements in performance in the primary validation study are observed, although the number of samples is small (Table 3). The areas under the ROC curves, shown in Figure 3B for the ≥6 month and ≥1 year secondary validation samples are 0.80 ± 0.14 and 0.86 ± 0.09, respectively.

Prevalent population studies

The validation study samples were highly enriched in rejection samples in order to more accurately estimate agreement with biopsy for this important but relatively rare class. In order to determine algorithm performance on the distribution of patients expected to be encountered in clinical practice, we tested 281 CARGO samples from 166 patients ≥1 year post-transplant, consisting of 160 (56.9%) grade 0, 68 (24.1%) grade 1A, 23 (8.1%) grade 1B, 21 (7.4%) grade 2 and 9 (3.2%) grade ≥3A, similar to that of the entire CARGO database. The grade 1B scores were significantly higher than grade 0 scores ($p = 0.0004$),

Table 2: Genes that discriminate between quiescence and rejection

Gene	GenBank	Discovery Source	PCR training *t*-test	PCR training Ratio
Programmed Cell death 1	PDCD1	Literature	1.6E–05	1.46
Semaphorin 7A	SEMA7A	Array	6.3E–05	1.29
Interleukin-1 receptor-soluble form	IL1R2	Array	6.4E–05	0.48
Importin alpha7	KPNA6	Array	1.1E–04	1.13
Chemokine (C-X-C motif) receptor 3	CXCR3	Literature	1.1E–04	1.32
Ikaros	ZNFN1A1	Literature	1.5E–04	1.15
Integrin beta 7	ITGB7	Literature	2.7E–04	1.21
Integrin alpha-M	ITGAM	Literature	3.6E–04	0.85
Chemokine (C-X-C motif) receptor 4	CXCR4	Array	4.6E–04	0.60
Matrix metalloproteinase 9	MMP9	Literature	6.1E–04	0.30
Vanin-2	VNN2	Array	6.4E–04	0.70
FLT3 ligand	FLT3LG	Literature	7.0E–04	1.27
CD160 NK cell receptor	BY55	Array	8.6E–04	1.21
Integrin alpha4	ITGA4	Literature	0.0011	1.18
Lymphocyte specific kinase	LCK	Literature	1.3E–03	1.27
T-cell receptor beta	TCRB	Literature	1.4E–03	1.30
Adenosine deaminase	ADA	Literature	0.0015	1.24
Adrenomedullin	ADM	Array	0.0016	0.70
Fms-like tyrosine kinase 3	FLT3	Literature	0.0020	0.60
Inositol polyphosphate-5-phosphatase	INPP5A	Array	0.0022	0.89
Lymphocyte activation gene 3	LAG3	Literature	0.0026	1.30
Fas Ligand	TNFSF6	Literature	0.0027	1.59
Signal regulatory protein beta-1	SIRPB1	Literature	0.0029	0.76
Carboxypeptidase M	CPM	Literature	0.0030	0.79
DAP12 associating lectin 1	CLECSF5	Literature	0.0030	0.72
Platelet factor 4	PF4	Literature	0.0032	0.74
Immune associated nucleotide receptor 7	hIAN7	Array	0.0032	1.26
Calgranulin A	S100A8	Array	0.0048	0.69
Guanine nucleotide exchange factor	VAV1	Literature	0.0057	1.08
Thrombopoietin receptor	MPL	Literature	0.0061	0.84
G6b Inhibitory receptor	G6b	Literature	0.0068	0.67
Ras homolog gene family, member U	ARHU	Array	0.0068	1.20
Notch homolog 1	NOTCH1	Literature	0.0073	1.11
Cas-Br-M (murine) ecotropic retroviral transforming sequence	CBLB	Literature	0.0081	1.15
T-cell transcription factor	GATA3	Literature	0.0095	1.15
T-cell R alpha	TCRA	Literature	0.0096	1.11
Calgranulin B	S100A9	Array*	0.0098	0.80
IL18	IL18	Literature	0.01	0.84
CD8A antigen	CD8A	Literature	0.01	1.13
Bruton's tyrosine kinase	BTK	Array	0.01	0.93
Granulysin	GNLY	Literature	0.02	1.29
CD28 antigen	CD28	Literature	0.02	1.13
Immunoglobulin J polypeptide	IGJ	Array*	0.02	1.49
erythrocyte membrane protein band 4.1	EPB41	Array*	0.02	1.18
Zn finger containing protein	ZFYVE27	Array	0.02	1.10
Cellular mediator of immune response	MIR	Array	0.03	1.14
Disabled homolog 1	DAB1	Array	0.03	0.93
Hemoglobin gamma	HBG	Array*	0.03	1.35
Integrin alpha2b	ITGA2B	Literature	0.03	0.64
CD10 antigen	CD10	Literature	0.04	0.68
Tumor necrosis factor receptor superfamily, member 7	TNFRSF7	Literature	0.04	1.24
PKC theta	PRKCQ	Literature	0.04	1.11
IgG heavy chain	IgHG	Array*	0.04	1.89
tumor necrosis factor (ligand) superfamily, member 4	TNFSF4	Literature	0.05	0.86
MIP-1-beta	CCL4	Array	0.07	1.27
Perforin	PRF1	Literature	0.08	1.32
OX40 receptor	TNFRSF4	Literature	0.08	1.32
Karyopherin beta 1	KPNB1	Array	0.10	1.01
WD40 motif bone marrow protein	WDR40A	Array	0.11	1.16
Tryptophanyl-tRNA synthetase	WARS	Array	0.14	1.27
Interleukin 8	IL8	Literature	0.20	1.62
Chemokine (C-X-C motif) ligand 10	CXCL10	Literature	0.22	1.35
Aminolevulinate, delta-, synthase 2	ALAS2	Array*	0.26	1.43
Chemokine (C-X-C motif) ligand 9	CXCL9	Literature	0.34	1.58
Hemoglobin alpha	HBA	Array*	0.36	1.36
Interleukin 7 receptor	IL7R	Literature	0.37	1.04
Major histocompatibility complex, class I, F	HLA-F	Array	0.42	1.27
Hemoglobin beta	HBB	Array*	0.62	1.32

*Correlated to significant array gene.

grade 1A scores ($p = 0.001$) and grade 2 scores ($p = 0.01$), but similar to grade ≥3A scores. The classifier scores and the prevalence of each biopsy grade in the CARGO database were used to estimate PPV and NPV. At a threshold of 30, the PPV is 6.8%, the NPV is 99.6%, and 68% of the tests are estimated to be below this value.

Discussion

This study tested the hypothesis that PBMC gene expression analysis using real time PCR for multiple genes and pathways detects the absence of moderate/severe acute cellular cardiac allograft rejection (ISHLT grade ≥3A) thus potentially reducing the frequency of graft biopsy in certain clinical settings. The multicenter validation study of independent patients showed that an 11 gene test distinguishes these states ($p = 0.0018$) and that patients ≥1 year

Table 3: Classifier performance in training, primary and secondary validation studies

			Biopsy grade 3A rejection				Biopsy grade 0 quiescence			
Sample set	Months post-Tx	Threshold	#Patients	#Samples	#Agree	%Agree	#Patients	#Samples	#Agree	%Agree
Training	All	20	29	36	29*	80.0%*	99	109	64*	59.0%*
1° Validation	All	20	31	31	26	83.9%	32	32	12	37.5%
2° Validation	All	20	50	62	47	75.8%	83	122	51	41.8%
1° Validation	>6	28	12	12	10	83.3%	14	14	10	71.4%
2° Validation	>6	28	19	21	15	71.4%	38	47	37	78.7%
1° Validation	>12	30	6	6	6	100%	7	7	4	57.1%
2° Validation	>12	30	10	10	8	80.0%	15	18	14	77.8%

*Bootstrap estimates.

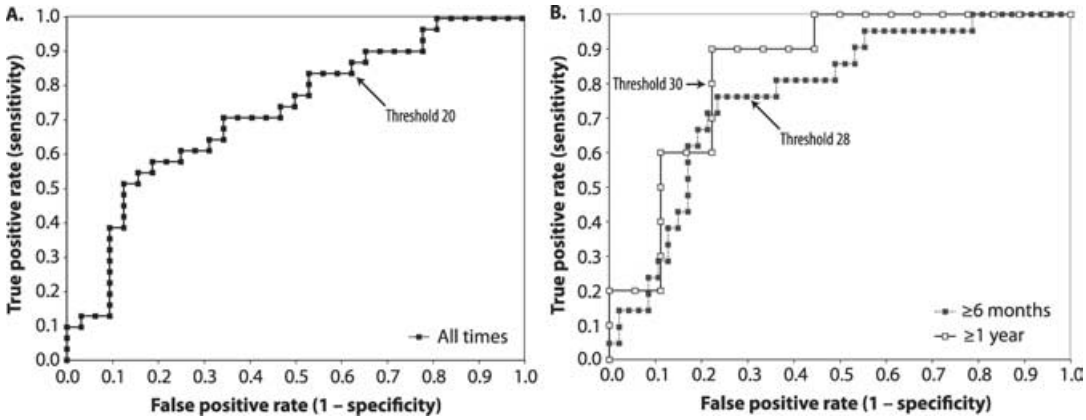


Figure 3: (A) Receiver operator characteristic (ROC) curve for primary validation cohort of 63 unique patient samples (31 rejection, 32 quiescent). Area under the ROC curve is 0.72 ± 0.06. **(B) ROC curves for secondary validation set of 184 patient samples (62 rejection, 122 quiescent) for the ≥6 months and ≥1 year periods.** Areas under the ROC curves are: 0.80 ± 0.114 for ≥6 months and 0.86 ± 0.09 for ≥1 year.

post-transplant with low molecular scores have a very low risk of current moderate/severe rejection (NPV >99%).

The specific genes identified as discriminating between quiescence (ISHLT grade 0) and moderate/severe rejection (ISHLT ≥ 3A) in PBMC-derived RNA encompass a wide variety of mechanisms and cell types (Table 2). Genes identified in T-cell mediated rejection in other organs have focused primarily on intragraft expression, with a more limited set of studies on peripheral fluids, including blood and urine. We found that both perforin and FasL/TNFSF6 as well as granulysin, were up-regulated in our study (see Table 2). All of these have been identified in renal allograft rejection studies by a variety of authors. In the study, which most closely resembles ours in terms of methodology and focus, PBMC gene expression was examined by RT-PCR of cytokine and cytokine receptor genes in heart allograft rejection (22). Of the four genes identified as most highly discriminatory in their study ($p < 0.01$), we also identified three as significant (CXCR3, FasL/TNFSF6 and perforin, Table 2), and we did not measure the other gene (COX2). Despite their significance, these genes were not included in the final LDA, because either they were not as robust as others or due to effects of sample processing on their expression. FasL/TNFSF6 and perforin cluster with PDCD1, the most discriminatory gene we identified (Table 2). PDCD1 has been identified as an important gene in an animal model of cardiac transplant rejection (34). CXCR3 and FasL/TNFSF6 were eliminated from consideration in the LDA, because expression of these genes is systematically dependent on sample processing time.

Although the primary endpoint of the study was achieved, important technical and clinical limitations of this study will have to be addressed to further evaluate the clinical role of this approach.

The derivation of the discriminant equation was critically dependent upon pathological classifications by multiple independent readers. The study design assumed a gold standard clinical endpoint of biopsy-based detection of rejection. However, the CARGO study demonstrated that this gold standard was limited by considerable inter-observer variability (30). Further work will be needed to assess how this variability is reflected in gene expression studies and clinical outcomes.

In the gene discovery phase, two approaches were taken: a focused leukocyte microarray and a knowledge-base or literature review, similar to that used to derive a validated

PCR-based test for breast cancer recurrence (35). Our goal was to find a set of genes that could be reproducibly measured by RT-PCR in a PBMC RNA preparation. The microarray approach used was limited to genes expressed in leukocytes potentially ignoring important genome-wide interactions. The knowledge-base approach, focusing on known genes, was ignoring new biology, which might be apparent in a nonhypothesis-driven approach. Although these complementary approaches led to the identification of genes that can distinguish the different rejection states, whole genome arrays may yield additional, different or better gene candidates. Moreover, additional studies may help to determine the basis for the differences observed between the microarray and PCR methods in significance of specific genes. These may be due to (i) lower sensitivity of microarrays leading to the elimination of genes which show discrimination in PCR, (ii) enhanced reproducibility of RT-PCR allowing measurement of small differences in gene expression (changes as small as 10–20%), likely undetectable by microarrays (usually eliminating genes that show <1.5- to 2-fold changes), (iii) use of a variety of classification definitions in the microarray analysis, which was focused to a single definition for the diagnostic development and validation phases.

Our microarray study results have to be interpreted in the context of recent insights from cancer biology-related array-based prediction studies (33). This reanalysis of seven studies showed that the list of genes identified as predictors of prognosis was highly unstable, and molecular signatures strongly depended on the selection of patients in the training sets. It highlights the primary challenge in using microarray results, where one is patient-limited and gene-rich: whether genes and signatures are truly significant or whether they are products of random variation (i.e. over-fitted to noise). We attempted to address two sources of variation—experimental (variability due to measurement technology) and biological (variability due to individual genetic and environmental differences). Experimental variation in microarrays was addressed by testing all candidate genes from the microarray discovery experiments with RT-PCR. Of the 97 candidate genes derived from microarray analyses, 27 were subsequently validated with RT-PCR (Table 2) and were used for further development. Biological variation was addressed in the RT-PCR data from the training set with statistical methods closely related to and more extensive than the suggested multiple random validation strategy (33), including multiple cross-validation and balanced bootstrapping techniques employed to calculate confidence intervals for performance measurements (details described in Supplement). Despite these results, it is still possible that some components of the LDA classifier are not robust; however, in using a two-technology, multiple-phase, cross-validated approach to validate discriminatory genes and the LDA signature, we have gone beyond most validation approaches previously described.

This study represents the first multicenter validation of a gene expression test in cardiac transplantation to identify patients at low risk for moderate/severe rejection. However, there are several limitations on the clinical implications of this work. Episodes of mild rejection on biopsy cannot be ruled out with this test. In addition, although the test captures patients with moderate/severe rejection, it has a low positive predictive value relative to biopsy. A nonquiescent score requires full workup including EMB to differentiate ongoing or impending cellular rejection, antibody-mediated rejection or chronic rejection/vasculopathy. Although CMV infection does not appear to confound the molecular signature (36,37), the impact of other infections on the test are not known. This study also did not address the predictive capacity of molecular testing for future rejection and clinical events, which could enable improved management in immunosuppressive therapy.

These study findings and limitations suggest several promising areas for future research including (A) more extensive gene discovery using whole-genome-based array approaches, (B) mechanistic hypothesis-driven research into individual genes and pathways identified and their intragraft role in acute rejection, (C) evaluation of the approach and specific genes and pathways in the setting of other solid organ transplantation. In summary, these results show that gene expression testing of blood cells can detect the absence of moderate/severe rejection, thus avoiding biopsy in certain clinical settings, but additional clinical experience is necessary to conclusively establish the predictive capacity of molecular testing for clinical events and its utility for monitoring immunosuppression.

Supplemental Material

The following supplemental material is available for this article online:

Appendix S1. Supplemental methods.
This material is available as part of the online article from http://www.blackwell-synergy.com

Acknowledgments

This study was supported in part by a grant from XDx, Inc. The authors wish to thank Drs Eric Topol (Cleveland Clinical Foundation), Jeff Bluestone (University of California, San Francisco), Ralph Snyderman (Duke University) and Fred Cohen (University of California, San Francisco) for their insightful reviews of the manuscript, Dr Thomas Quertermous (Stanford) for advice on scientific issues, and Drs Rob Tibshirani and Mike Walker (Stanford) for invaluable support and advice on statistical design and analysis and Dr Adriana Zeevi (Pittsburgh) for contributions to the study and review of the manuscript. The study was made possible through the hard work of the many CARGO investigators and researchers listed in the supplemental material.

References

1. Eisen HJ, Tuzcu EM, Dorent R et al. Everolimus for the prevention of allograft rejection and vasculopathy in cardiac-transplant recipients. N Engl J Med 2003; 349: 847–858.
2. Kobashigawa J, Miller L, Renlund D et al. A randomized active-controlled trial of mycophenolate mofetil in heart transplant recipients. Mycophenolate Mofetil Investigators. Transplantation 1998; 66: 507–515.
3. Nielsen H, Sorensen FB, Nielsen B, Bagger JP, Thayssen P, Baandrup U. Reproducibility of the acute rejection diagnosis in human cardiac allografts. The Stanford Classification and the International Grading System. J Heart Lung Transplant 1993; 12: 239–243.
4. Winters GL, McManus B. Consistencies and controversies in the application of the International Society for Heart and Lung Transplantation working formulation for heart transplant biopsy specimens. Rapamycin Cardiac Rejection Treatment Trial Pathologists. J Heart Lung Transplant 1996; 15: 728–735.
5. Winters GL, Marboe CC, Billingham M. The International Society for Heart and Lung Transplantation grading system for heart transplant biopsy specimen: clarification and commentary. J Heart Lung Transplant 1998; 17: 754–760.
6. Angermann CE, Nassau K, Stempfle HU et al. Recognition of acute cardiac allograft rejection from serial integrated backscatter analyses in human orthotopic heart transplant recipients. Comparison with conventional echocardiography. Circulation 1997; 95: 140–150.
7. Bourge R, Eisen H, Hershberger R et al. Noninvasive rejection monitoring of cardiac transplants using high resolution intramyocardial electrograms: initial US multicenter experience. PACE 1998; 21: 2338–44.
8. Deng MC, Erren M, Roeder N et al. T-cell and monocyte subsets, inflammatory molecules, rejection and hemodynamics early after cardiac transplantation. Transplantation 1998; 65: 1255–1261.
9. Eisen HJ, Eisenberg S, Saffitz JE, Bolman RM, Sobel BE, Bergmann S. Noninvasive diagnosis of cardiac transplant rejection using labeled lymphocytes. Circulation 1987; 75: 868.
10. Frist W, Yasuda T, Segall G et al. Noninvasive detection of human cardiac transplant rejection with indium-111 antimyosin (Fab) imaging. Circulation 1987; 76: 81–85.
11. Hammer C, Klanke D, Lersch C et al. Cyoimmunologic monitoring (CIM) for differentiation between cardiac rejection and viral, bacterial, or fungal infection: its specificity and sensitivity. Transplant Proc 1989; 21: 3631–3633.
12. Hummel M, Dandel M, Knollmann F et al. Long-term surveillance of heart-transplanted patients: noninvasive monitoring of acute rejection episodes and transplant vasculopathy. Transplant Proc 2001; 33: 3539–3542.
13. Kimball P, Radovancevic B, Isom T, Frazier B, Kerman R. Cytokine panel predicts early rejection and therapeutic response after cardiac transplantation. Transplant Proc 1995; 27: 1286–1287.
14. Knosalla C, Grauhan O, Muller J et al. Intramyocardial electrogram recordings (IMEG) for diagnosis of cellular and humoral mediated cardiac allograft rejection. Ann Thorac Cardiovasc Surg 2000; 6: 89–94.
15. Masuyama T, Valantine HA, Gibbons R, Schnittger I, Popp R. Serial measurement of integrated ultrasonic backscatter in human cardiac allografts forthe recognition of acute rejection. Circulation 1990; 81: 829–839.
16. Melter M, Exeni A, Reinders ME et al. Expression of the chemokine receptor CXCR3 and its ligand IP-10 during human cardiac allograft rejection. Circulation 2001; 104: 2558–2564.
17. Park JW, Warnecke H, Deng M, Schueler S, Heinrich KW, Hetzer R. Early diastolic left ventricular function as a marker of acute cardiac rejection: a prospective serial echocardiographic study. Int J Cardiol 1992; 37: 351–359.
18. Pfitzmann R, Muller J, Grauhan O, Hetzer R. Intramyocardial impedance measurements for diagnosis of acute cardiac allograft rejection. Ann Thorac Surg 2000; 70: 527–532.
19. Smart FW, Young JB, Weilbaecher D, Kleiman NS, Wendt RE 3rd, Johnston D. Magnetic resonance imaging for assessment of tissue rejection after heterotopic heart transplantation. J Heart Lung Transplant 1993; 12: 403–410.
20. Warnecke H, Muller J, Cohnert T et al. Clinical heart transplantation without routine endomyocardial biopsy. J Heart Lung Transplant 1992; 11: 1093–1102.
21. Horwitz PA, Tsai EJ, Putt ME et al. Detection of cardiac allograft rejection and response to immunosuppressive therapy with peripheral blood gene expression. Circulation 2004; 110: 3815–3821.
22. Schoels M, Dengler TJ, Richter R, Meuer SC, Giese T. Detection of cardiac allograft rejection by real-time PCR analysis of circulating mononuclear cells. Clin Transplant 2004; 18: 513–517.
23. Walter SD, Irwig L, Glasziou P. Meta-analysis of diagnostic tests with imperfect reference standards. J Clin Epidemiol 1999; 52: 943–951.
24. Zhou X, Obuchowski N, McClish D. Statistical Methods in Diagnostic Medicine. 2002, New York: John Wiley & Sons.
25. GEO: http://www.ncbi.nlm.nih.gov/geo/
26. Brunet JP, Tamayo P, Golub TR, Mesirov J. Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci U S A 2004; 101: 4164–4169.
27. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 2002; 99: 6567–6572.
28. www.unos.org
29. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A, 2001; 98: 5116–5121.
30. Marboe CC, Billingham M, Eisen H et al. Nodular endocardial infiltrates (Quilty Lesions) cause significant variability in the diagnosis of ISHLT rejection grades 2 and 3A in endomyocardial biopsies from cardiac allograft recipients. J Heart Lung Transplant 2005; 24: S219–S226.
31. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998; 95: 14863–14868.
32. Baechler EC, Batliwalla FM, Karypis G et al. Expression levels for many genes in human peripheral blood cells are highly sensitive to ex vivo incubation. Genes Immunology 2004; 5: 347–353.
33. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. Lancet 2005; 365: 488–492.
34. Ozkaynak E, Wang L, Goodearl A et al. Programmed death-1 targeting can promote allograft survival. J Immunol 2002; 169: 6546–6553.
35. Paik S, Shak S, Tang G et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004; 351: 2817–2826.
36. Deng MC, Mehra M, Hunt S et al. Leukocyte gene expression signature of CMV viremia in cardiac allograft recipients is distinct from that seen during allograft rejection. Am J Transplant, 2004. 4e: 455.
37. Jenner RG, Young R. Insights into host responses against pathogens from transcriptional profiling. Nat Rev Microbiol 2005; 3: 281–294.

Molecular Testing for Long-term Rejection Surveillance in Heart Transplant Recipients: Design of the Invasive Monitoring Attenuation Through Gene Expression (IMAGE) Trial

Michael X. Pham, MD, MPH,[a,b] Mario C. Deng, MD,[c] Abdallah G. Kfoury, MD,[d] Jeffrey J. Teuteberg, MD,[e] Randall C. Starling, MD, MPH,[f] and Hannah Valantine, MD[a]

Background: Acute rejection continues to occur beyond the first year after cardiac transplantation, but the optimal strategy for detecting rejection during this late period is still controversial. Gene expression profiling (GEP), with its high negative predictive value for acute cellular rejection (ACR), appears to be well suited to identify low-risk patients who can be safely managed without routine invasive endomyocardial biopsy (EMB).

Methods: The Invasive Monitoring Attenuation Through Gene Expression (IMAGE) study is a prospective, multicenter, non-blinded, randomized clinical trial designed to test the hypothesis that a primarily non-invasive rejection surveillance strategy utilizing GEP testing is not inferior to an invasive EMB-based strategy with respect to cardiac allograft dysfunction, rejection with hemodynamic compromise (HDC) and all-cause mortality.

Results: A total of 199 heart transplant recipients in their second through fifth post-transplant years have been enrolled in the IMAGE study since January 13, 2005. The study is expected to continue through 2008.

Conclusions: The IMAGE study is the first randomized, controlled comparison of two rejection surveillance strategies measuring outcomes in heart transplant recipients who are beyond their first year post-transplant. The move away from routine histologic evaluation for allograft rejection represents an important paradigm shift in cardiac transplantation, and the results of this study have important implications for the future management of heart transplant patients. J Heart Lung Transplant 2007; 26:808–14. Copyright © 2007 by the International Society for Heart and Lung Transplantation.

Success after cardiac transplantation is still limited by acute cardiac allograft rejection. Even with contemporary immunosuppression, 30% to 50% of patients will have acute cellular rejection (ACR) within the first year after transplantation.[1,2] Although the incidence of ACR declines dramatically after the first year,[3,4] several observational studies have reported that clinically significant ACR continues to occur in this time period.[5–8] Although many cases of late rejection resolve without augmentation of immunosuppression and carry a benign initial prognosis,[5] studies have suggested that the number of ACR episodes within the first 2 years after transplantation is predictive of long-term survival.[9] Rejection episodes requiring augmented immunosuppression, even when occurring beyond the first post-transplant year, have also been shown to increase the risk of developing cardiac allograft vasculopathy (CAV).[10] Thus, many transplant centers continue to perform routine rejection surveillance beyond the first year after transplantation.

The EMB is currently the standard for identifying acute rejection[11] yet is associated with rare but potentially serious risks, including vascular injury,[12] tricuspid valve damage,[13] ventricular perforation and tamponade, arrhythmias[14] and heart block.[15] In addition, EMB is costly,[16] uncomfortable for patients, and subject to both sampling error and inter-observer variability.[17] These limitations underscore the need for a non-invasive monitoring strategy that would safely decrease the number of EMBs performed, particularly in the follow-up of patients after 1 year post-transplantation, when the risk of acute rejection is lower.[18,19]

From the [a]Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California; [b]Cardiology Section, VA Palo Alto Health Care System, Palo Alto, California, [c]Department of Medicine, Columbia University, New York, New York; [d]LDS Hospital, Salt Lake City, Utah; [e]Cardiovascular Institute, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; and [f]Department of Cardiovascular Medicine, Cleveland Clinic, Cleveland, Ohio.

Submitted March 21, 2007; revised May 9, 2007; accepted May 28, 2007.

Supported by XDx, Inc.

Reprint requests: Michael X. Pham, MD, Division of Cardiovascular Medicine, Stanford University School of Medicine, Falk CVRB, 300 Pasteur Drive, Stanford, CA 94305. Telephone: 650-723-6141. Fax: 650-725-1599. E-mail: mpham@cvmed.stanford.edu

 1053-2498/07/$–see front matter. doi:10.1016/j.healun.2007.05.017

Gene expression profiling (GEP) is based on the analysis of peripheral blood mononuclear cell ribonucleic acid (RNA) using real-time polymerase chain reaction (PCR) technology. Previous studies have shown that certain gene expression profiles correlate with histologic rejection on EMB, and their analysis may provide a valuable tool for the non-invasive monitoring of acute rejection.[20,21] In the multicenter Cardiac Allograft Rejection Gene Expression Observation (CARGO) study, a multigene algorithm based on the expression of 20 genes (11 informative, 9 controls) was developed and validated. The informative genes are involved in T-cell activation and trafficking, natural killer (NK)-cell activation, stem cell mobilization, hematopoiesis and alloimmune recognition.[22] The algorithm weighs the contribution of each gene and results in a score ranging from 0 and 40, with scores below threshold indicating a very low likelihood of moderate-to-severe ACR on EMB (ISHLT Grade ≥3A/2R). In the CARGO study, GEP scores below 30, 32 and 34 were associated with negative predictive values of 99.6%, 99.4% and 99.2%, respectively, when used to monitor stable outpatients who were beyond the first year post-transplant.[22,23] Based on these findings, the GEP test has been used clinically at many transplant centers since January 2005 as part of their rejection surveillance protocols.[23]

Studies of GEP testing to date have focused on correlation with EMB results and prediction of future rejection events, but no studies have directly compared patient outcomes between rejection surveillance protocols primarily employing GEP testing vs traditional protocols based on routine EMB. We therefore undertook this study to evaluate the impact of a non-invasive rejection surveillance strategy employing GEP, as compared with an invasive strategy utilizing routine EMB, on graft dysfunction, death and hemodynamically compromising (HDC) acute rejection. We hypothesized that a GEP-based strategy would result in equivalent clinical outcomes while decreasing the need for biopsies, lowering resource utilization, and improving patient quality of life (QOL).

METHODS

Study Objectives

The Invasive Monitoring Attenuation Through Gene Expression (IMAGE) study is a prospective, multi-center, non-blinded, randomized clinical trial designed to test the hypothesis that a non-invasive rejection surveillance strategy utilizing GEP is not inferior to a traditional, invasive strategy utilizing routine EMB, with respect to clinical outcomes (decrease in left ventricular ejection fraction [LVEF], acute rejection with associated HDC, or death) when used to monitor asymptomatic heart transplant patients. Our secondary objectives are to determine whether there are any differences in biopsy-related complications, QOL and health-care resource utilization between the two strategies.

Patient Population

The IMAGE study expects to enroll approximately 505 adult (≥18 years) heart transplant recipients who are in their second to fifth year post-transplant (>12 months and ≤60 months) and are seen for routine rejection surveillance. Inclusion criteria include: (a) clinically stable cardiac allograft function, defined as an LVEF of ≥45% by echocardiography, multiple gated acquisition (MUGA) scan or ventriculography at the time of study enrollment; and (b) absence of prior or current evidence of either severe cardiac allograft vasculopathy (CAV) or antibody-mediated rejection (AMR) with associated HDC. Key exclusion criteria include presence of signs or symptoms of cardiac allograft dysfunction, rejection therapy for ISHLT Grade 3A/2R or higher rejection during the 60 days preceding enrollment, use of ≥20 mg/day of prednisone or equivalent doses of corticosteroids at the time of study enrollment, and end-stage renal disease requiring long-term renal replacement therapy. A full list of eligibility criteria are presented in Table 1. Approval from the institutional review board (IRB) of each participating institution was obtained prior to patient enrollment.

End-points

The primary end-point is a composite of: (a) ≥25% proportional decrease in LVEF on echocardiography relative to the enrollment value; (b) any rejection with HDC; and (c) all-cause mortality. HDC is defined by the presence of one or more of the following: an absolute LVEF ≤30% or a drop in LVEF ≥25% compared with the baseline value; a cardiac index <2.0 liters/min/m^2; or the use of inotropic drugs to support circulation.

Secondary and exploratory end-points include: number of biopsies performed; number of biopsy-related complications; health-care resource utilization; and QOL associated with each rejection surveillance strategy.

Study Procedures

After screening and obtaining written informed consent, patients are randomized in a 1:1 fashion to either the GEP- or EMB-based arm. Randomization is stratified by study center and by year post-transplant (second and third vs fourth and fifth post-transplant year) to account for differential follow-up of patients in these groups and to ensure balance of these factors among the two study arms. Patients are followed for 24 months after enrollment or until a total of 54 primary events are reached in the study, whichever occurs first. All patients, regardless of randomization status, undergo rejection surveil-

Table 1. IMAGE Elibility Criteria

Inclusion criteria:

1. Heart transplant recipients in the second to fifth year (>12–60 months) post-transplant.
2. Age ≥18 years.
3. Stable outpatient being seen for routine monitoring of rejection. Stability is defined as absence of prior or current evidence of either severe cardiac allograft vasculopathy (CAV) or antibody-mediated rejection (AMR) with associated hemodynamic compromise.
   (a) Severe CAV is defined as either: (A) >50% left main stenosis; (B) ≥50% stenosis in ≥2 primary vessels (proximal one third or middle one third of the LAD or LCx, RCA to take-off of PDA in right-dominant coronary circulations); or (C) isolated branch stenoses of >50% in all three systems (diagonal branches, obtuse marginal branches, distal one third of LAD or LCx artery, PDA, PLB, and RCA to takeoff of PDA in non-dominant systems).
   (b) AMR with associated hemodynamic compromise is defined as AMR (defined according to local criteria) with either: (A) a left ventricular ejection fraction (LVEF) ≤30% or ≥25% lower than the baseline value; (B) a cardiac index <2 liters/min/m^2; or (C) the use of inotropic agents to support circulation.
4. Left ventricular ejection fraction ≥45% by echocardiography, multiple gated acquisition (MUGA) scan, or ventriculography at study entry (baseline/enrollment study).

Exclusion criteria:

1. Patients <1 year after heart transplantation.
2. Any symptoms or clinical signs of impaired allograft function:
   (a) Symptoms of decompensated heart failure at the enrollment visit.
   (b) Signs of decompensated heart failure, including the development of a new S3 gallop at the enrollment visit.
   (c) Elevated right heart pressures with diminished cardiac index <2.2 liters/min/m^2 that is new compared to a previous measurement within 6 months.
   (d) Decrease in LVEF as measured by echocardiography: ≥25% compared with prior measurement within 6 months.
3. Rejection therapy for biopsy-proven ISHLT Grade 3A or higher during the preceding 2 months.
4. Major changes in immunosuppression therapy within previous 30 days (e.g., discontinuation of calcineurin inhibitors, switch from mycophenolate mofetil to sirolimus or vice versa).
5. Unable to give written informed consent.
6. Patient receiving hematopoietic growth factors (e.g., Neupogen, Epogen) currently or during the previous 30 days.
7. Patients receiving ≥20 mg/day of prednisone-equivalent corticosteroids at the time of enrollment.
8. Patient enrolled in a trial requiring routine surveillance endomyocardial biopsies.
9. Patient received transfusion within preceding 4 weeks.
10. Patients with end-stage renal disease requiring some form of renal replacement therapy (hemodialysis or peritoneal dialysis).
11. Pregnancy at the time of enrollment.

LAD, left anterior descending; LCx, left circumflex; RCA, right coronary artery; PDA, posterior descending artery; PLB, posterolateral branch.

lance at pre-specified time intervals ranging from 3 to 12 months, based on their transplant center protocol and time post-transplant. At the time of each rejection surveillance visit, patients undergo a full history and physical examination, assessment of cardiac allograft function by echocardiography, and monitoring for cardiac allograft rejection using either GEP or EMB. Patients complete a QOL and health questionnaire at enrollment, after 1 year of follow-up, and at the completion of the study.

Procedures Specific to EMB Group

Patients in the EMB group are monitored for rejection using a combination of clinical assessment, echocardiography, and EMB, according to each center's existing protocols. In addition, they undergo GEP testing at the time of each study visit, but the treating physicians are blinded to the test results. GEP scores in this group will be evaluated retrospectively to determine whether a positive GEP score and negative biopsy is predictive of subsequent events such as graft failure, HDC acute rejection or death.

Procedures Specific to GEP Group

Patients in the GEP group undergo GEP testing either immediately prior to or at the time of each rejection surveillance visit. Patients with clinical or echocardiographic evidence of cardiac allograft dysfunction also undergo EMB and further diagnostic testing (i.e., coronary angiography) according to center-specific protocols. Treatment decisions are based on the results of the EMB and other conventional diagnostic testing. If the EMB shows treatable rejection or if the clinician decides to initiate anti-rejection treatment with pulse steroids, antibody therapy or apheresis, then patients are monitored using center-specific protocols and conventional rejection surveillance methods (typically EMB) for 3 months after the rejection episode. Patients are returned to the GEP arm after 3 event-free months, defined as absence of rejection requiring treatment, heart failure

symptoms, or graft dysfunction by echocardiography according to the study end-point definition.

In the event that symptoms of graft dysfunction or rejection occur between follow-up clinic visits, the patient returns immediately for re-assessment of graft function by echocardiography, and an EMB may be performed at the discretion of the treating physician to evaluate for rejection.

Patients with clinical and echocardiographic evidence of stable allograft function are further managed based on the results of their GEP test. Patients with a GEP score of <34 are considered to be at very low risk for moderate/severe ACR and are treated as if they had no evidence of rejection. Patients with a GEP score of ≥ 34 return for EMB within 5 days after the initial clinic visit and are managed according to the EMB result. If the EMB does not show evidence of treatable rejection, anti-rejection therapy is withheld, assuming there is no graft dysfunction. If the EMB demonstrates treatable rejection as determined by the individual center, patients are treated according to center-specific protocols and are returned to the GEP arm after 3 event-free months.

Patients with consistently high GEP scores above threshold (score ≥ 34) on three consecutive visits and who have no clinical manifestations of graft dysfunction or treatable rejection on two consecutive EMBs over a period of 3 to 9 months may be considered for management without a surveillance EMB on the third visit if there is no statistically significant increase in GEP score. Two scores are considered statistically different from one another if the previous score does not fall within

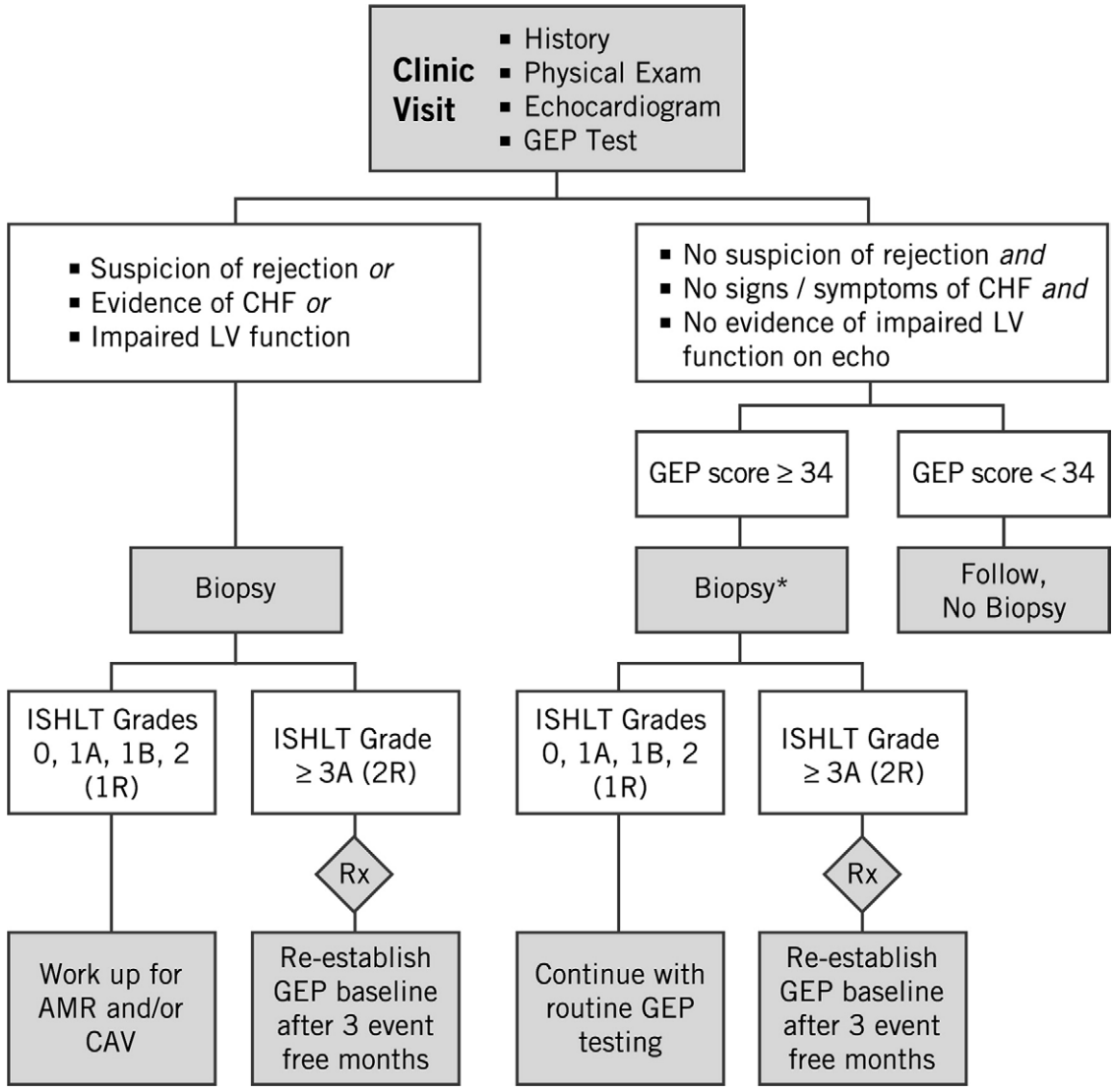


Figure 1. Study schema for patients enrolled in the GEP surveillance arm. *Consider no biopsy if ≥ 3 GEP scores ≥ 34 with absence of treatable rejection on previous two biopsies. CHF, congestive heart failure; ISHLT, International Society for Heart and Lung Transplantation; AMR, antibody-mediated rejection; CAV, cardiac allograft vasculopathy.

the 95% confidence interval of the current score. The study scheme for the GEP group is summarized in Figure 1.

GEP Testing

GEP testing is performed using the AlloMap molecular expression test developed by XDx, Inc. (Brisbane, CA). The AlloMap test uses the same 20-gene assay and classifier developed in the CARGO study.[22] Approximately 8 ml of venous blood is collected by routine phlebotomy. The sample is processed locally and shipped frozen to the XDx laboratory for testing. An AlloMap score, accompanied by a 95% confidence interval, is reported to the transplant center within 5 calendar days.

Statistical Methods

The primary analysis will be a non-inferiority analysis comparing overall event-free survival times, in an intention-to-treat manner, between the two study arms. Kaplan–Meier plots and summary statistics will be used to describe the event-free survival in each group, and the Cox proportional hazards model will be used to compare the two arms with respect to the primary end-point. A hazard ratio and associated 95% confidence interval will serve as the basis for non-inferiority. The protocol-defined non-inferiority margin for the hazard ratio confidence interval is set at 2.054. Assuming a primary event rate of 10% per year in both arms and maximum follow-up of 2 years, it is estimated that approximately 505 patients will be required to accrue 54 events. This will provide 80% power to test the hypothesis that a GEP-based rejection surveillance protocol is non-inferior to an EMB-based approach with respect to the primary end-point, at a 1-sided level of significance of 5%.

Interim Analysis

An independent data safety monitoring board (DSMB) will review the efficacy and safety data every 6 months to determine if the study should be stopped early, continued or terminated. In addition, an interim analysis comparing the safety data and event-free survival between each study arm is planned when 50% of the primary end-points have occurred across both study arms. Pocock boundaries for early termination of the study are calculated using the Lan–DeMets method. The Pocock boundaries assign an $\alpha = 0.0310$ for the planned interim analysis, and an $\alpha = 0.0297$ for the final analysis, while still preserving an overall 5% Type I error rate. If the confidence interval falls within the non-inferiority margin at the interim analysis, the trial will be stopped. If it does not, there will be a conditional power analysis assuming that the observed event rates will apply for the remainder of the target sample size. If the conditional power is $<50\%$, the trial will be stopped for futility.[24] The study will otherwise continue until the target number of end-points is achieved.

QOL and Economic Analyses

Health-care utilization for patients in each arm will be estimated using data on the number of biopsies and other major procedures performed as well as length of stay and discharge diagnoses associated with transplant-related hospitalizations. QOL will be evaluated using utility measures and a standardized health survey (Short Form-12) that measures 8 domains of health: physical functioning; role limitations due to physical health; bodily pain; general health perceptions; vitality; social functioning; role limitations due to emotional problems; and mental health.

Study Organization

IMAGE currently involves six participating transplant centers in the USA. We anticipate that two or three additional centers will be added in 2007. The study is sponsored by XDx, Inc. A scientific steering committee, which includes the principal investigators from each of the participating centers, meets regularly to oversee and provide scientific direction for the study. An independent DSMB conducts efficacy and safety data review every 6 months, in addition to a planned interim. An end-points committee will review and adjudicate all cases of acute rejection with HDC. Finally, a blinded core echocardiography laboratory is responsible for measurement of LVEF. A full listing of IMAGE investigators, committee members and clinical sites is presented in the Appendix.

RESULTS

A total of 199 patients have been enrolled in the IMAGE study since January 13, 2005. The study is expected to continue through 2008 with an interim analysis expected in 2007.

DISCUSSION

The IMAGE study compares a non-invasive, GEP-based rejection surveillance protocol to a traditional invasive (EMB-based) protocol, for monitoring heart transplant recipients in their second to fifth years post-transplant. The study will evaluate the impact of these two strategies on the incidence of cardiac allograft dysfunction, development of acute rejection with HDC, and all-cause mortality.

Selection of End-points

We selected a composite primary end-point to include events that have an equal chance of being detected in both arms, assuming comparable efficacy of both rejec-

tion surveillance techniques. In addition to all-cause mortality, we chose ACR with HDC and left ventricular systolic dysfunction because asymptomatic rejection could presumably progress to overt heart failure symptoms, and/or graft dysfunction in the absence of timely diagnosis and anti-rejection therapy. Furthermore, ACR with associated HDC has been shown to be associated with poor short-term patient survival.[25]

Rationale for Identifying an Appropriate GEP Threshold

A GEP threshold of ≥30 was initially used to identify patients at risk for moderate/severe acute cellular rejection. This conservative threshold was chosen based on the high negative predictive value (NPV) at this cut-off for patients in the CARGO study who were beyond 1 year post-transplantation.[22] On November 7, 2005, the protocol was amended to reflect a change in the GEP threshold to ≥34. The decision to increase the GEP threshold was made after a review of GEP performance characteristics estimated from the CARGO population across a range of scoring thresholds and time post-transplant.[23] Changing the threshold to ≥34 resulted in a negligible change in the NPV of the test (99.6% to 99.2%), but it decreased the percent of scores above threshold from 50.8% to 22.3% and resulted in fewer biopsies performed due to high (above threshold) GEP scores.

In conclusion, the IMAGE study is the first randomized, controlled comparison of two rejection surveillance strategies in heart transplantation. The use of early markers of immune activation to monitor for allograft rejection represents a novel strategy in solid-organ transplantation. In addition, the move away from routine histologic surveillance for rejection represents an important evolution in heart transplantation, historically following both kidney and liver transplantation, where "protocol" biopsies are rarely performed and most biopsies are event-driven. The results of this study have important implications for the future management of heart transplant recipients.

REFERENCES

1. Hershberger RE, Starling RC, Eisen HJ, et al. Daclizumab to prevent rejection after cardiac transplantation. N Engl J Med 2005;352:2705–13.
2. Eisen HJ, Tuzcu EM, Dorent R, et al. Everolimus for the prevention of allograft rejection and vasculopathy in cardiac-transplant recipients. N Engl J Med 2003;349:847–58.
3. White JA, Guiraudon C, Pflugfelder PW, Kostuk WJ. Routine surveillance myocardial biopsies are unnecessary beyond one year after heart transplantation. J Heart Lung Transplant 1995;14:1052–6.
4. Sethi GK, Kosaraju S, Arabia FA, Rosado LJ, McCarthy MS, Copeland JG. Is it necessary to perform surveillance endomyocardial biopsies in heart transplant recipients? J Heart Lung Transplant 1995;14:1047–51.
5. Klingenberg R, Koch A, Schnabel PA, et al. Allograft rejection of ISHLT grade >/=3A occurring late after heart transplantation—a distinct entity? J Heart Lung Transplant 2003;22:1005–13.
6. Gradek WQ, D'Amico C, Smith AL, Vega D, Book WM. Routine surveillance endomyocardial biopsy continues to detect significant rejection late after heart transplantation. J Heart Lung Transplant 2001;20:497–502.
7. Winters GL, Costanzo-Nordin MR, O'Sullivan EJ, et al. Predictors of late acute orthotopic heart transplant rejection. Circulation 1989;80:III-106–110.
8. Stehlik J, Starling RC, Movsesian MA, et al. Utility of long-term surveillance endomyocardial biopsy: a multi-institutional analysis. J Heart Lung Transplant 2006;25:1402–9.
9. Radovancevic B, Konuralp C, Vrtovec B, et al. Factors predicting 10-year survival after heart transplantation. J Heart Lung Transplant 2005;24:156–9.
10. Brunner-La Rocca HP, Schneider J, Kunzli A, Turina M, Kiowski W. Cardiac allograft rejection late after transplantation is a risk factor for graft coronary artery disease. Transplantation 1998;65:538–43.
11. Caves PK, Stinson EB, Billingham M, Shumway NE. Percutaneous transvenous endomyocardial biopsy in human heart recipients. Experience with a new technique. Ann Thorac Surg 1973;16:325–36.
12. Baraldi-Junkins C, Levin HR, Kasper EK, Rayburn BK, Herskowitz A, Baughman KL. Complications of endomyocardial biopsy in heart transplant patients. J Heart Lung Transplant 1993;12:63–7.
13. Williams MJ, Lee MY, DiSalvo TG, et al. Biopsy-induced flail tricuspid leaflet and tricuspid regurgitation following orthotopic cardiac transplantation. Am J Cardiol 1996;77:1339–44.
14. Oldham N, Ott RA, Allen BA, Fopiano P, Dwyer M. Ventricular fibrillation complicating endomyocardial biopsy of a cardiac allograft. Cathet Cardiovasc Diagn 1991;23:300–1.
15. Cui G, Kobashigawa J, Margarian A, Sen L. Cause of atrioventricular block in patients after heart transplantation. Transplantation 2003;76:137–42.
16. Evans RW, Williams GE, Baron HM, et al. The economic implications of noninvasive molecular testing for cardiac allograft rejection. Am J Transplant 2005;5:1553–8.
17. Marboe CC, Billingham M, Eisen H, et al. Nodular endocardial infiltrates (Quilty lesions) cause significant variability in diagnosis of ISHLT Grade 2 and 3A rejection in cardiac allograft recipients. J Heart Lung Transplant 2005;24(suppl):S219–226.
18. Kubo SH, Naftel DC, Mills RM Jr, et al. Risk factors for late recurrent rejection after heart transplantation: a multiinstitutional, multivariable analysis. Cardiac Transplant Research Database Group. J Heart Lung Transplant 1995;14:409–18.
19. Kirklin JK, Naftel DC, Bourge RC, et al. Rejection after cardiac transplantation. A time-related risk factor analysis. Circulation 1992;86(suppl):II-236–41.
20. Horwitz PA, Tsai EJ, Putt ME, et al. Detection of cardiac allograft rejection and response to immunosuppressive therapy with peripheral blood gene expression. Circulation 2004;110:3815–21.
21. Schoels M, Dengler TJ, Richter R, Meuer SC, Giese T. Detection of cardiac allograft rejection by real-time PCR analysis of circulating mononuclear cells. Clin Transplant 2004;18:513–7.
22. Deng MC, Eisen HJ, Mehra MR, et al. Noninvasive discrimination of rejection in cardiac allograft recipients using gene expression profiling. Am J Transplant 2006;6:150–60.
23. Starling RC, Pham M, Valantine H, et al. Molecular testing in the management of cardiac transplant recipients: initial clinical experience. J Heart Lung Transplant 2006;25:1389–95.
24. Chen YH, DeMets DL, Lan KK. Increasing the sample size when the unblinded interim result is promising. Stat Med 2004;23:1023–38.

25. Mills RM, Naftel DC, Kirklin JK, et al. Heart transplant rejection with hemodynamic compromise: a multiinstitutional study of the role of endomyocardial cellular infiltrate. Cardiac Transplant Research Database. J Heart Lung Transplant 1997;16:813–21.

APPENDIX

IMAGE Investigators and Sites

Steering Committee: Hannah Valantine, MD, MRCP (Study Chair), Michael Pham, MD, MPH (Study Co-Chair), Allen Anderson, MD, Mario C. Deng, MD, A. G. Kfoury, MD, Randall C. Starling, MD, MPH, Jeffrey J. Teuteberg, MD, Helen Baron, MD, Farhad Kazazi, PhD.

Data Safety Monitoring Board: Robert Bourge, MD (Chair), Maryl Johnson, MD, David Naftel, PhD, Si Pham, MD.

End-points Committee: Roberta Bogaev, MD, Michael Felker, MD, Barbara Pisani, MD, Lynne Wagoner, MD.

Core Echocardiography Laboratory: David Liang, MD, PhD (Director), Stanford University Medical Center.

Clinical Sites: Stanford University Medical Center and VA Palo Alto Health Care System, Palo Alto, CA; New York Presbyterian Hospital, New York, NY; LDS Hospital, Salt Lake City, UT; University of Pittsburgh Medical Center, Pittsburgh, PA; Cleveland Clinic, Cleveland, OH, University of Chicago Medical Center, Chicago, IL.

ORIGINAL ARTICLE

Gene-Expression Profiling for Rejection Surveillance after Cardiac Transplantation

Michael X. Pham, M.D., M.P.H., Jeffrey J. Teuteberg, M.D., Abdallah G. Kfoury, M.D., Randall C. Starling, M.D., M.P.H., Mario C. Deng, M.D., Thomas P. Cappola, M.D., Andrew Kao, M.D., Allen S. Anderson, M.D., William G. Cotts, M.D., Gregory A. Ewald, M.D., David A. Baran, M.D., Roberta C. Bogaev, M.D., Barbara Elashoff, M.S., Helen Baron, M.D., James Yee, M.D., Ph.D., and Hannah A. Valantine, M.D., for the IMAGE Study Group*

ABSTRACT

BACKGROUND

Endomyocardial biopsy is the standard method of monitoring for rejection in recipients of a cardiac transplant. However, this procedure is uncomfortable, and there are risks associated with it. Gene-expression profiling of peripheral-blood specimens has been shown to correlate with the results of an endomyocardial biopsy.

METHODS

We randomly assigned 602 patients who had undergone cardiac transplantation 6 months to 5 years previously to be monitored for rejection with the use of gene-expression profiling or with the use of routine endomyocardial biopsies, in addition to clinical and echocardiographic assessment of graft function. We performed a noninferiority comparison of the two approaches with respect to the composite primary outcome of rejection with hemodynamic compromise, graft dysfunction due to other causes, death, or retransplantation.

RESULTS

During a median follow-up period of 19 months, patients who were monitored with gene-expression profiling and those who underwent routine biopsies had similar 2-year cumulative rates of the composite primary outcome (14.5% and 15.3%, respectively; hazard ratio with gene-expression profiling, 1.04; 95% confidence interval, 0.67 to 1.68). The 2-year rates of death from any cause were also similar in the two groups (6.3% and 5.5%, respectively; $P=0.82$). Patients who were monitored with the use of gene-expression profiling underwent fewer biopsies per person-year of follow-up than did patients who were monitored with the use of endomyocardial biopsies (0.5 vs. 3.0, $P<0.001$).

CONCLUSIONS

Among selected patients who had received a cardiac transplant more than 6 months previously and who were at a low risk for rejection, a strategy of monitoring for rejection that involved gene-expression profiling, as compared with routine biopsies, was not associated with an increased risk of serious adverse outcomes and resulted in the performance of significantly fewer biopsies. (ClinicalTrials.gov number, NCT00351559.)

From Stanford University Medical Center, Stanford (M.X.P., H.A.V.), and VA Palo Alto Health Care System, Palo Alto (M.X.P.) — both in California; the University of Pittsburgh Medical Center, Pittsburgh (J.J.T.); Intermountain Medical Center and Intermountain Healthcare, Salt Lake City (A.G.K.); Cleveland Clinic, Cleveland (R.C.S.); Columbia University Medical Center, New York (M.C.D.); Hospital of the University of Pennsylvania, Philadelphia (T.P.C.); Mid America Heart Institute, Saint Luke's Hospital, Kansas City, MO (A.K.); University of Chicago Medical Center (A.S.A.) and Northwestern University (W.G.C.) — both in Chicago; Washington University School of Medicine, St. Louis (G.A.E.); Newark Beth Israel Medical Center, Newark, NJ (D.A.B.); Texas Heart Institute, Houston (R.C.B.); and XDx, Brisbane, CA (B.E., H.B., J.Y.). Address reprint requests to Dr. Valantine at the Division of Cardiovascular Medicine, Stanford University School of Medicine, Falk CVRB, 300 Pasteur Dr., Stanford, CA 94305, or at hvalantine@stanford.edu.

*The members of the Invasive Monitoring Attenuation through Gene Expression (IMAGE) Study Group are listed in the Appendix.

This article (10.1056/NEJMoa0912965) was published on April 22, 2010, at NEJM.org.

N Engl J Med 2010;362:1890-900.

Advances in immunosuppression after cardiac transplantation have increased the rates of 1-year survival among recipients to nearly 90%. However, acute cellular rejection is still observed during the first year after transplantation (at a rate of approximately 30 to 40%) and occurs at a lower rate thereafter.[1-4] Rejection episodes are associated with an increased risk of allograft vasculopathy and loss.[5-7] Endomyocardial biopsy has remained the primary method of monitoring for rejection, despite the discomfort and the rare but potentially serious complications of the procedure.[8-12]

Quantitative assessment of mononuclear-cell gene expression in peripheral-blood specimens has been explored as a method for detecting the rejection of a cardiac transplant.[13,14] This approach has been investigated as an alternative to an endomyocardial biopsy[13,14] and has led to the development and validation of a commercially available test that has been shown to correlate with the results of an endomyocardial biopsy.[14] Although this gene-expression test has been used at some cardiac transplantation centers to monitor transplant recipients for rejection,[15] it has not been compared systematically in clinical practice with the current standard approach to monitoring for rejection with the use of routine biopsies.

We conducted the Invasive Monitoring Attenuation through Gene Expression (IMAGE) trial to test the hypothesis that a strategy of monitoring for rejection that involves gene-expression profiling is not inferior to a strategy that involves routine biopsies, with respect to a composite outcome of rejection with hemodynamic compromise, graft dysfunction, death, or retransplantation.

Methods

Study Design and Oversight

The IMAGE study was a randomized, event-driven, noninferiority trial conducted at 13 U.S. cardiac transplantation centers from January 2005 through October 2009. The study design has been described previously,[16] and additional details are included in the Supplementary Appendix, available with the full text of this article at NEJM.org. The trial was sponsored by XDx, in which Stanford University owns equity; XDx is the maker of the AlloMap test. The academic investigators initiated and designed the study in collaboration with the sponsor. The trial protocol was approved by the institutional review board at each participating center. The sponsor was involved in the collection and source verification of the data, and the sponsor's biostatisticians performed the analyses with oversight from the study steering committee. The first author wrote the initial draft of the manuscript, and revisions were made by all the authors. Investigators at the core echocardiography laboratory at Stanford University reread all the echocardiograms to calculate the left ventricular ejection fractions that were used in the analyses. An independent end-points committee adjudicated all primary events. A data and safety monitoring board monitored efficacy and safety data. The academic investigators vouch for the accuracy and completeness of the data and of all analyses.

Patients

Patients 18 years of age or older who had undergone a cardiac transplantation between 1 and 5 years previously were eligible for enrollment. Data on cardiac transplantations at participating centers were obtained from the Organ Procurement and Transplantation Network of the United Network of Organ Sharing (http://optn.transplant.hrsa.gov). A protocol amendment on November 27, 2007, expanded enrollment to include patients who had undergone a cardiac transplantation more than 6 months previously, in order to facilitate enrollment. At the time of enrollment, patients were required to be in a clinically stable condition and to have a left ventricular ejection fraction of 45% or greater. Exclusion criteria included a history of severe allograft vasculopathy, antibody-mediated rejection, or the presence of signs or symptoms of heart failure. All participating patients provided written informed consent.

Study Procedures

Patients were randomly assigned, in a 1:1 ratio, to undergo monitoring for rejection by means of gene-expression profiling (gene-profiling group) or routine endomyocardial biopsies (biopsy group). Randomization was stratified according to study center and according to the interval since transplantation (1 year or less, 2 to 3 years, or 4 to 5 years). Monitoring for rejection with the use of the assigned strategy was performed at prespecified intervals in both groups according to the

protocols at the individual transplantation centers (see Table 2 in the Supplementary Appendix). All patients in both groups were also monitored with the use of clinical and echocardiographic assessments. The performance of a biopsy was mandated by the protocol for patients in both groups if clinical or echocardiographic evidence of graft dysfunction was present or, in the case of the gene-profiling group, if the gene-expression profiling score was above a specified threshold. If patients had consistently elevated gene-expression profiling scores and no evidence of rejection on at least two previous biopsies, the protocol did not require further biopsies to be performed in the case of a third or subsequent instance of a score above the threshold.

Gene-expression testing was performed with the use of the AlloMap test (XDx), which evaluates expression levels of 11 informative genes that were shown in previous studies to distinguish between rejection and the absence of rejection. Possible scores range from 0 to 40, with higher scores having a stronger correlation with histologic rejection. In a previous study, a score below 30 had a negative predictive value of 99.6% for histologic evidence of rejection.[14] Therefore, the initial protocol for the current trial specified a score of 30 as a threshold for a mandatory biopsy. However, on November 7, 2005, the protocol was amended to increase the threshold for a mandatory biopsy to a score of 34 in order to minimize the number of biopsies that would be needed in the gene-profiling group. Further details of the test and of the characteristics of the test threshold are provided in the Supplementary Appendix.

Patients were followed for a maximum of 24 months, until they died, or until the study completion date, whichever occurred first. The study ended in October 2009, after the minimum prespecified number of primary outcome events (54) had occurred.

OUTCOMES

The primary outcome was the first occurrence of rejection with hemodynamic compromise, graft dysfunction due to other causes, death, or retransplantation. Definitions of each component of the composite primary outcome are provided in the Supplementary Appendix. Secondary outcomes included death from any cause, the number of biopsies performed, and biopsy-related complications. We also assessed the patients' quality of life and their satisfaction with the method of monitoring for rejection. Quality of life was assessed with the use of the Medical Outcomes Study 12-Item Short Form Health Survey (SF-12). We assessed satisfaction by asking the patients, "How satisfied are you with the current method of detecting rejection?" Responses were scored on an ordinal scale that ranged from 1 (very unhappy) to 10 (very happy).

STATISTICAL ANALYSIS

The trial was designed to test for the noninferiority of gene-expression profiling, as compared with routine endomyocardial biopsies, with respect to the primary outcome. The primary analysis, which was conducted in the intention-to-treat population, was a comparison between the groups of the time to the first occurrence of the composite primary outcome; the comparison was made with the use of the hazard ratios calculated from a Cox proportional-hazards model. The strategy of gene-expression profiling was considered to be noninferior to the strategy of routine biopsies if the one-sided upper boundary of the 95% confidence interval for the hazard ratio with the gene-expression–profiling strategy, as compared with the biopsy strategy, was less than the prespecified margin for noninferiority (2.054). This relative margin was derived assuming a primary-event rate of 5% per year in the biopsy group and the possibility of an event rate of up to 10% per year in the gene-profiling group. This difference, in the view of the investigators, would balance the expected benefit with respect to patient convenience and satisfaction that would result from a reduction in the number of biopsies performed. The study required that a minimum of 54 primary events occur, in order to exclude the inferiority null hypothesis with 80% power, assuming an overall event rate of 5% per year, a rate that was estimated from published observational data.[17,18]

Means and standard deviations for continuous variables were calculated and compared with the use of Student's t-test. Numbers and proportions for categorical variables were compared with the use of Fisher's exact test. Both the Kaplan–Meier method and Cox proportional-hazards models were used to estimate event rates. The effects of an interaction between strategy group and center and between strategy group and interval between transplantation and randomization were tested at an alpha level of 0.15.

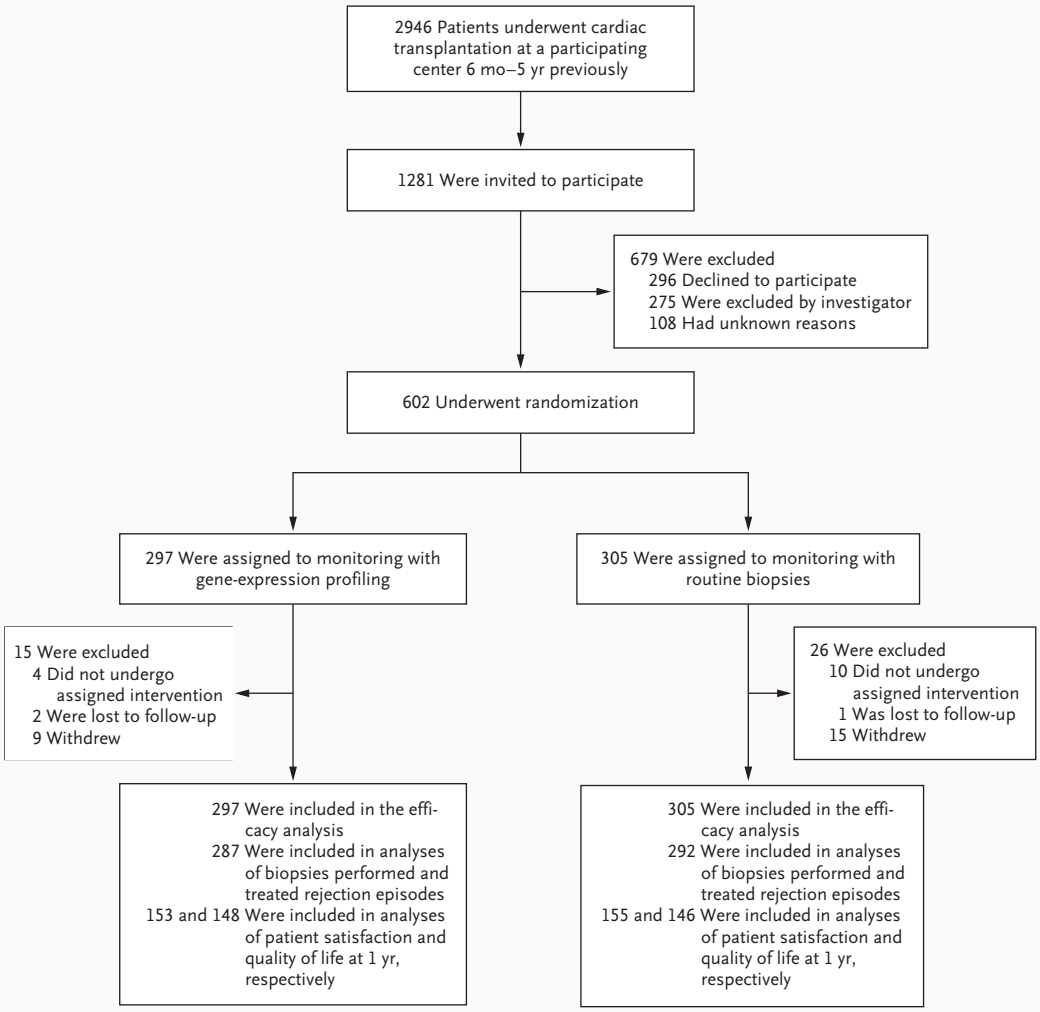


Figure 1. Screening, Randomization, and Inclusion in Analyses.

During the study enrollment period (2005 through 2009), there were 2946 adults who had undergone cardiac transplantation at a participating center between 6 months and 5 years previously, as documented by the Organ Procurement and Transplantation Network of the United Network of Organ Sharing (http://optn.transplant.hrsa.gov). Participating centers were asked to screen all potentially eligible patients for enrollment in the study. A total of 1665 of the 2946 potentially eligible patients (57%) were either not approached for consent or did not meet the eligibility criteria at the time of screening. Details regarding the patients who did not meet the eligibility criteria are not available. The reason that an investigator elected not to enroll a patient was not routinely recorded; however, investigators were encouraged to preferentially enroll patients who were in the early post-transplantation period (<3 years), since data for these patients were expected to be most meaningful. The two other most common reasons for an investigator electing not to enroll a patient were a complicated medical course and the preference of the treating physician to continue with biopsy-based monitoring for rejection. The analyses of biopsies performed and treated rejection episodes included data from patients who completed at least one study visit and who were followed for a minimum of 30 days in the study. Both scheduled study visits and unscheduled outpatient visits were included.

RESULTS

PATIENTS

A total of 602 patients were randomly assigned to be monitored for rejection with the use of gene-expression profiling or with the use of routine endomyocardial biopsies (Fig. 1). The baseline characteristics of the two groups were well matched except that there was a higher proportion of black patients in the biopsy group than in the gene-

profiling group (P=0.01) (Table 1). The interval between transplantation and randomization was 6 through 12 months in the case of 15% of the patients, 13 through 36 months in the case of 68%, and 37 through 60 months in the case of 17%. The median duration of follow-up after randomization was 19.0 months (interquartile range, 9.6 to 23.8).

Table 1. Baseline Characteristics of the Study Population.*

Characteristic	Gene Profiling (N=297)	Biopsy (N=305)	P Value
Age — yr			
Mean	53.9±12.9	54.3±12.8	0.68
Range	18.0–74.0	19.0–78.0	
Male sex — no. (%)	244 (82.2)	249 (81.6)	0.92
Race or ethnic group — no. (%)†			
White	236 (79.5)	232 (76.1)	0.33
Hispanic	22 (7.4)	17 (5.6)	0.41
Black	25 (8.4)	46 (15.1)	0.01
Asian or Pacific Islander	7 (2.4)	6 (2.0)	0.79
Other	7 (2.4)	4 (1.3)	0.38
Indication for cardiac transplantation — no. (%)			0.96
Coronary artery disease	127 (42.8)	130 (42.6)	
Nonischemic cardiomyopathy	152 (51.2)	155 (50.8)	
Valvular heart disease	6 (2.0)	5 (1.6)	
Congenital heart disease	9 (3.0)	9 (3.0)	
Graft vasculopathy or retransplantation	1 (0.3)	3 (1.0)	
Other	2 (0.7)	3 (1.0)	
Interval between transplantation and randomization — no. (%)			
6–12 mo	43 (14.5)	44 (14.4)	>0.99
13–36 mo	205 (69.0)	208 (68.2)	0.86
37–60 mo	49 (16.5)	53 (17.4)	0.83
Cytomegalovirus status — no. (%)			
Donor and recipient positive	128 (43.1)	109 (35.7)	0.07
Donor and recipient negative	44 (14.8)	47 (15.4)	0.91
Donor positive and recipient negative	59 (19.9)	78 (25.6)	0.10
Donor negative and recipient positive	50 (16.8)	58 (19.0)	0.52
Unknown	16 (5.4)	13 (4.3)	
Use of ventricular assist device before transplantation — no. (%)	58 (19.5)	57 (18.7)	0.84
Induction therapy — no. (%)			
Any	168 (56.6)	181 (59.3)	0.74
Muromonab-CD3	4 (1.3)	5 (1.6)	
Antithymocyte globulin	52 (17.5)	53 (17.4)	
Basiliximab	30 (10.1)	43 (14.1)	
Daclizumab	66 (22.2)	63 (20.7)	
Alemtuzumab	12 (4.0)	13 (4.3)	
Other	4 (1.3)	4 (1.3)	

Table 1. (Continued.)

Characteristic	Gene Profiling (N=297)	Biopsy (N=305)	P Value
Immunosuppressive therapy — no. (%)‡			
Cyclosporine	79 (26.6)	83 (27.2)	0.66
Tacrolimus	218 (73.4)	218 (71.5)	0.65
Mycophenolate mofetil or mycophenolic acid	237 (79.8)	250 (82.0)	0.53
Azathioprine	26 (8.8)	15 (4.9)	0.08
Sirolimus	53 (17.8)	65 (21.3)	0.31
Prednisone	132 (44.4)	122 (40.0)	0.28
Medical history after transplantation — no. (%)			
Hypertension treated with medication	247 (83.2)	258 (84.6)	0.66
Diabetes mellitus treated with medication	115 (38.7)	114 (37.4)	0.74
Renal insufficiency§	147 (49.5)	157 (51.5)	0.68
Lipid-lowering drug prescribed	275 (92.6)	283 (92.8)	>0.99
Cancer	38 (12.8)	49 (16.1)	0.30
Left ventricular ejection fraction at first study visit¶	63.2±6.0	63.4±6.1	0.67

* Plus–minus values are means ±SD. Data are for the intention-to-treat population.
† Race or ethnic group was self-reported.
‡ This category includes all medications taken by patients while they were enrolled in the study.
§ Renal insufficiency was defined by a serum creatinine level of less than 1.5 mg per deciliter (133 μmol per liter).
¶ Data for first-visit measurements of left ventricular ejection fraction were missing for 9 patients in the gene-profiling group and 15 in the biopsy group. In the case of five patients in the gene-profiling group and seven in the biopsy group, the left ventricular ejection fraction could not be calculated at the core echocardiography laboratory owing to the poor quality of the echocardiogram. In these cases, the measurement of left ventricular ejection fraction that was obtained closest to the first study visit was used in the analysis.

PRIMARY OUTCOME

The 2-year rate of the composite primary outcome in the gene-profiling group was similar to the rate in the biopsy group (14.5% and 15.3%, respectively; P=0.86) (Fig. 2A). The corresponding hazard ratio was 1.04 (95% confidence interval [CI], 0.67 to 1.68), with the upper boundary falling below the prespecified noninferiority margin. Therefore, monitoring for rejection with gene-expression profiling was noninferior to monitoring with routine biopsies with respect to the prevention of the primary outcome. The results for the individual components of the primary outcome are shown in Table 2. There was no significant interaction with respect to the primary outcome between the assigned group and either the interval between transplantation and randomization (≤12 months vs. >12 months) or the transplantation center (P=0.86 and P=0.99, respectively). Because there was a higher proportion of black patients in the biopsy group than in the gene-profiling group and a higher observed rate of the primary outcome among black patients than among nonblack patients (18.3% vs. 10.2%, P=0.07), we performed an additional analysis adjusting for black race in our Cox model. We found that the results were consistent with those of our primary analysis (hazard ratio, 1.13; 95% CI, 0.70 to 1.84).

DEATHS

The overall rate of survival in our study did not differ significantly according to the method of monitoring (Fig. 2B). The 2-year cumulative rate of death was 6.3% in the gene-profiling group and 5.5% in the biopsy group (P=0.82) (Table 2). The adjudicated causes of death (cardiovascular vs. noncardiovascular) were similar in the groups.

BIOPSIES PERFORMED AND RELATED COMPLICATIONS

A total of 409 biopsies were performed in the gene-profiling group, as compared with 1249 performed in the biopsy group (Fig. 3, and Table 8 in the Supplementary Appendix). The frequency of

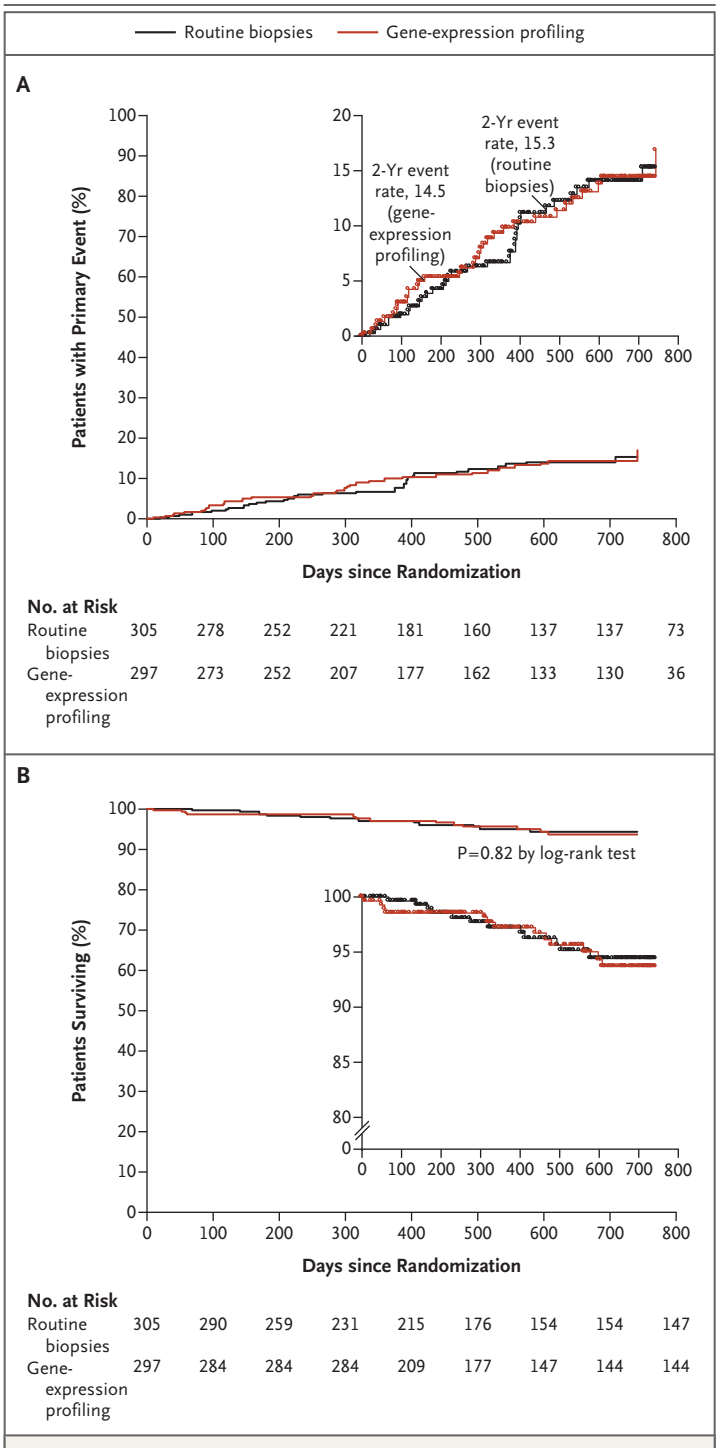


Figure 2. Kaplan–Meier Estimates of the Time to the Composite Primary Outcome and the Probability of Survival.

Panel A shows the time to the first occurrence of any of the following primary events: rejection with hemodynamic compromise, graft dysfunction due to other causes, death, or retransplantation. Only the first event that was part of the composite primary outcome was considered. Panel B shows the probability of overall survival. The inset in each panel shows the same data on an enlarged y axis and on a condensed x axis.

biopsies was 0.5 biopsies per patient-year of follow-up in the gene-profiling group and 3.0 biopsies per patient-year of follow-up in the biopsy group (P<0.001). In the gene-profiling group, 67% of the biopsies were performed because of elevated gene-expression profiling scores; another 17% were performed, per protocol, when signs, symptoms, or echocardiographic manifestations of graft dysfunction were present at the time of a clinic visit, 13% were performed as part of a follow-up assessment after treatment for rejection, and 3% were performed outside the study protocol. In 28 instances (9% of the cases in which there were elevated scores), consistently high gene-expression profiling scores did not result in performance of a biopsy (see the Methods section). Biopsy-related complications occurred in four patients in the biopsy group and in one patient assigned to the gene-profiling group (Table 2).

INTENSITY AND COMPLICATIONS OF IMMUNOSUPPRESSION

The overall intensity of immunosuppression throughout the study was similar in the gene-profiling group and the biopsy group (see the Supplementary Appendix). Despite a higher number of infections among patients monitored with gene-expression profiling than among those monitored with biopsies (53 vs. 43) (Table 7 in the Supplementary Appendix), we found no significant differences between the groups in mean levels of calcineurin inhibitors (Section 2.2 in the Supplementary Appendix). The mean serum creatinine levels during the study were also similar in the two groups (1.42±0.41 mg per deciliter [125.5±36.2 μmol per liter] in the gene-profiling group vs. 1.42±0.59 mg per deciliter [125.5±52.2 μmol per liter] in the biopsy group, P=0.95). Finally, the incidence of any cancer was similar in the two groups (3.7% in the gene-profiling group and 3.3% in the biopsy group, P=0.83).

REJECTION EPISODES

A total of 34 treated episodes of rejection occurred in the gene-profiling group, as compared with 47 episodes in the biopsy group (Section 2.4 and Table 9 in the Supplementary Appendix). In the gene-profiling group, six treated episodes of rejection were initially detected as a result of a biopsy performed because of an elevated gene-expression score. In the biopsy group, 22 treated episodes of rejection were asymptomatic and were detected on routine biopsy alone.

Table 2. Trial Outcomes.

Outcome	Total Events		2-Yr Cumulative Event Rate		P Value	Hazard Ratio (95% CI)*
	Gene Profiling	Biopsy	Gene Profiling	Biopsy		
			%			
Composite primary outcome — no. of events†	34‡	33	14.5	15.3	0.86§	1.04 (0.67–1.68)
Rejection with hemodynamic compromise as first event — no. of events	11	13			>0.99¶	
Cellular, biopsy-confirmed‖	2	7				
Antibody-mediated, biopsy-confirmed**	3	1				
Mixed, biopsy-confirmed	3	2				
Probable, not biopsy-confirmed††	4	3				
Graft dysfunction due to other causes as first event — no. of events	11	14			0.68¶	
Allograft vasculopathy	1	1				
Nonspecific graft failure	11	13				
Death as first event — no.	11	6			0.23¶	
Cardiovascular	7	5				
Noncardiovascular or unknown	4	1				
Death at any time — no. of events‡‡	13	12	6.3	5.5	0.82§	1.10 (0.50–2.40)
Cardiovascular	8	9				
Noncardiovascular or unknown	5	3				
Adverse events associated with biopsy — no. of patients/total no. (%)§§	1/287 (0.3)	4/292 (1.4)				
Tricuspid-valve incompetence¶¶	0/287	2/292				
Symptomatic pericardial effusion‖‖	0/287	1/292				
Bleeding	0/287	1/292				
Other***	1/287	0/292				

* The hazard ratio was estimated with the use of the Cox model, which included study-group assignment as a factor.

† The composite primary outcome was rejection with hemodynamic compromise, graft dysfunction due to other causes, death, or retransplantation. The analysis was performed on the basis of the first occurrence of any of the components. One patient in the biopsy group underwent retransplantation. The event was categorized as a death in the analyses, as specified by the statistical-analysis plan.

‡ One case of graft dysfunction was adjudicated as due to probable rejection (not biopsy-confirmed), allograft vasculopathy, or both. Therefore, this event is listed twice, once in the category of probable rejection and once in the category of allograft vasculopathy.

§ P values were calculated with the use of the log-rank test.

¶ P values were calculated with the use of Fisher's exact test for categorical variables.

‖ Confirmation of cellular rejection on biopsy required that a local pathologist classify the biopsy specimen, according to the International Society for Heart and Lung Transplantation system for grading rejection, as a grade of 2R (according to the 2004 version, in which the grades range from 0 to 3R) or 3A (according to the 1990 version, in which grades range from 0 to 4). Higher numbers indicate more severe rejection.

** Confirmation of antibody-mediated rejection on biopsy required histologic evidence of acute capillary injury or immunopathological evidence of antibody-mediated injury (as assessed with the use of immunofluorescence or immunohistochemical testing).

†† Probable rejection included events that, in the absence of histologic confirmation on biopsy, were considered by the end-points committee to be caused by rejection.

‡‡ This category includes deaths that occurred as the first event, as well as deaths that occurred after a nonfatal primary event.

§§ The total number includes all patients who completed at least one study visit and who were followed for a minimum of 30 days in the study.

¶¶ This category included tricuspid regurgitation with a grade of moderate-to-severe or higher on the basis of the local echocardiography report. One patient had severe tricuspid regurgitation, and the other patient had moderate-to-severe tricuspid regurgitation.

‖‖ Hypotension and presyncope developed in one patient in the biopsy group after a routine heart biopsy; an echocardiogram in this patient showed a new, moderate-grade pericardial effusion that required hospitalization for observation. The effusion resolved without the need for drainage.

*** One patient was inadvertently given subcutaneous formalin instead of lidocaine before venous cannulation, and the wound required local débridement by a plastic surgeon.

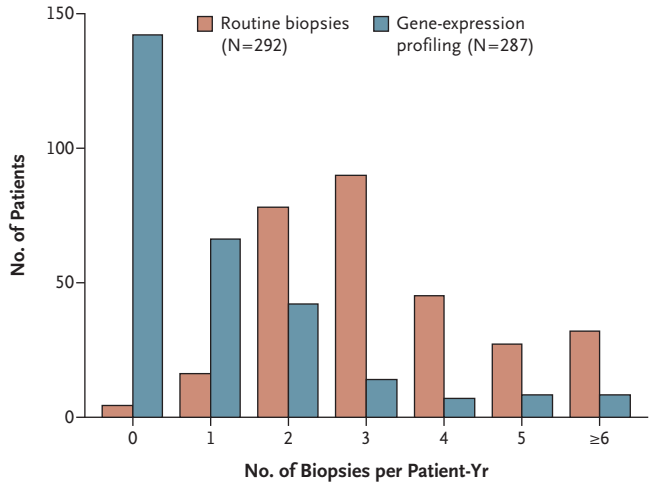


Figure 3. Frequency of Endomyocardial Biopsies Performed.

The distribution of outpatient endomyocardial biopsies performed per patient-year of follow-up is shown for patients in the biopsy group and patients in the gene-profiling group. The majority of patients in the gene-profiling group (88%) underwent two biopsies or fewer per patient-year, and 50% did not require a biopsy during the study.

HEALTH STATUS AND PATIENT SATISFACTION

At enrollment, no significant differences were found between the two groups in the physical-health and mental-health summary scores of the SF-12 (Table 10 in the Supplementary Appendix). The physical-health summary score was higher in the biopsy group than in the gene-profiling group at 1 year (47.3 vs. 44.7, P=0.03), but both the mean physical-health and mental-health summary scores were similar in the two groups at 2 years (physical-health score: 45.1 in the gene-profiling group and 46.2 in the biopsy group, P=0.52; mental-health score: 50.8 and 50.7 in the two groups, respectively; P=0.66). At enrollment, the scores for patient satisfaction were similar in the gene-profiling group and the biopsy group (6.86 and 6.74, respectively; P=0.61). During the course of the study, there was an increase in the satisfaction score in the gene-profiling group, to 8.15 in year 1 and 8.74 in year 2, whereas the scores in the biopsy group remained similar throughout the study to the score at enrollment (6.64 in year 1 and 6.66 in year 2). The differences in patient-satisfaction scores at 1 and 2 years between patients in the gene-profiling group and those in the biopsy group were significant (P<0.001 for both comparisons).

DISCUSSION

In this multicenter study involving patients who had received a cardiac transplant more than 6 months before enrollment and whose condition was clinically stable, the use of gene-expression profiling of peripheral-blood specimens in combination with clinical and echocardiographic assessment, as compared with the use of endomyocardial biopsies according to standard practice, resulted in a significant reduction in the number of biopsies performed and did not result in an excess of adverse outcomes. In addition, patient satisfaction was higher with the gene-expression profiling method of monitoring than with the biopsy method, reflecting the preference of many patients for avoiding an invasive procedure.

For gene-expression testing, a score below 34 was used in the majority of cases (97%) to identify patients who were at low risk for rejection and in whom a biopsy was not needed. Although the use of a higher threshold may further minimize the number of biopsies needed, the results of our trial suggest that a score below 34 represents a prudent threshold to use in clinical practice in the case of patients for whom the interval after transplantation is more than 6 months.

There were fewer treated episodes of rejection in the gene-profiling group than in the biopsy group, and this difference was due primarily to fewer asymptomatic episodes of rejection observed in the gene-profiling group than in the biopsy group (see the Supplementary Appendix). Although gene-expression profiling may not have detected all the cases of asymptomatic rejection, we did not observe an excess 2-year cumulative risk of graft dysfunction, death, or retransplantation in the gene-profiling group. This observation suggests that not all asymptomatic episodes of rejection that occur more than 6 months after transplantation warrant treatment. Some of these episodes may be explained by a misreading on the part of pathologists of benign nodular endocardial infiltrates (Quilty lesions) in biopsy specimens, whereas others may represent a subgroup of histologically defined rejection episodes that resolve without augmentation of immunosuppression.[19,20] Conversely, undetected rejection may lead to long-term graft dysfunction through such mechanisms as progressive myocardial fibrosis or coronary-artery intimal hyperplasia. The late consequences of untreated rejection are poorly understood and may not have been clinically ap-

parent during the follow-up period (a median of 19 months) in our study.

Only 6 of the 34 treated episodes of rejection in the gene-profiling group were detected with the use of the gene-expression test. The other episodes were detected because of the presence of overt symptoms of heart failure or echocardiographic evidence of graft dysfunction. These observations raise the possibility that clinical observation may detect the majority of serious rejection episodes. Some transplantation centers in the United States and many centers worldwide have discontinued the practice of performing routine biopsies after the first year post-transplantation.[21] However, many physicians who treat transplant recipients have been reluctant to adopt this practice until the relative safety of such an approach can be shown in a comparative trial. Therefore, the majority of transplantation centers in the United States continue to perform biopsies beyond the first year post-transplantation, although there is considerable institutional variation in the frequency and duration of monitoring for rejection.[2] There has not been sufficient equipoise to justify a comparison of monitoring by means of clinical observation with monitoring by means of routine biopsies, but our findings may provide the basis for such comparisons in future studies.

The results of our trial must be interpreted in the context of several important limitations. Only patients who had received a cardiac transplant more than 6 months previously were eligible for enrollment. Such patients have a lower risk of rejection and may be at lower risk for adverse outcomes due to undetected rejection than patients for whom the interval after transplantation is 6 months or less. We chose to enroll patients who were at a lower risk for rejection because the relative safety of an approach that minimizes the number of biopsies has not been confirmed, and we did not want to expose the study participants to an undue risk of adverse events. This decision reflects the characteristically conservative approach to the care of cardiac-transplant recipients and the reluctance of clinicians and patients to accept even a small possibility of causing harm.

Only 20% of potentially eligible patients were enrolled in the study. Patients who had received a cardiac transplant less than 3 years previously were recruited preferentially, and a substantial number of eligible patients were not enrolled, owing to patient or physician preferences. Details regarding these patients were not available, but it is likely that patient selection was biased toward the inclusion of low-risk patients, thereby restricting the generalizability of our findings. This limitation should be taken into account by clinicians when they consider the use of gene-expression profiling in the care of their patients.

The low projected event rates and the limited number of available patients necessitated the choice of a wide noninferiority margin. The trial's reduced power was reflected in a relatively wide confidence interval that does not exclude the possibility of a 33% decrease in primary event rates (or 1.8 fewer events per 100 patient-years) or of a 68% increase (3.7 excess events per 100 patient-years) among patients in the gene-profiling group. Our composite outcome was chosen to include both clinically overt rejection and the possible consequences of undiagnosed rejection. Because graft dysfunction, death, or retransplantation may be caused by conditions other than rejection, the inclusion of these end points may have further reduced the trial's power. A more robust test of noninferiority would have necessitated a considerably larger sample than that which was feasible, given the limited number of cardiac transplantations performed worldwide.[22]

Finally, the lack of blinding in the study may have influenced the intensity of immunosuppression in the gene-profiling group. However, we did not observe any significant differences between the groups in mean levels of calcineurin inhibitors throughout the study, in serum creatinine levels, or in the incidence of neoplasms.

In conclusion, our study suggests that gene-expression profiling of peripheral-blood specimens may offer a reasonable alternative to routine biopsies, for monitoring cardiac-transplant recipients for rejection if the interval since transplantation is at least 6 months and the patient is considered to be at low risk for rejection. However, the study had limited power to allow for a firm conclusion to be reached regarding the use of gene-expression profiling as a substitute for the performance of biopsies. A larger trial with a narrower noninferiority margin and a longer follow-up period would be necessary to definitively resolve this issue.

Supported by XDx, maker of the AlloMap test. Stanford University holds equity in XDx.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

APPENDIX

Members of the IMAGE Study Group are as follows: **Steering Committee:** H. Valantine (chair), M. Pham (co-chair), A. Anderson, D. Baran, R. Bogaev, T. Cappola, W. Cotts, M. Deng, G. Ewald, A. Kao, A. Kfoury, R. Starling, J. Teuteberg. **Data and Safety Monitoring Board:** R. Bourge (chair), M. Johnson, D. Naftel, S. Pham. **Endpoints Committee:** B. Edwards, M. Felker, L. Wagoner. **Echocardiography Core Laboratory:** D. Liang (director), J. Chow, A. Paloma, J. Puryear, A. Rodriguez (cardiac sonographers). **Data Coordinating Center:** G. Alexander, B. Elashoff, T. Klingler, S. Wang, H. Wolters (statisticians), U. Patil, J. Scheel (data management), A. Clark, D. Pieretti, E. Shocron, N. Sternheim, T. Wolf (clinical operations), H. Baron, J. Yee (clinical sciences).

REFERENCES

1. Hershberger RE, Starling RC, Eisen HJ, et al. Daclizumab to prevent rejection after cardiac transplantation. N Engl J Med 2005;352:2705-13.

2. Stehlik J, Starling RC, Movsesian MA, et al. Utility of long-term surveillance endomyocardial biopsy: a multi-institutional analysis. J Heart Lung Transplant 2006; 25:1402-9.

3. Gradek WQ, D'Amico C, Smith AL, Vega D, Book WM. Routine surveillance endomyocardial biopsy continues to detect significant rejection late after heart transplantation. J Heart Lung Transplant 2001;20:497-502. [Erratum, J Heart Lung Transplant 2001;20:803.]

4. Heimansohn DA, Robison RJ, Paris JM III, Matheny RG, Bogdon J, Shaar CJ. Routine surveillance endomyocardial biopsy: late rejection after heart transplantation. Ann Thorac Surg 1997;64:1231-6.

5. Radovancevic B, Konuralp C, Vrtovec B, et al. Factors predicting 10-year survival after heart transplantation. J Heart Lung Transplant 2005;24:156-9.

6. Raichlin E, Edwards BS, Kremers WK, et al. Acute cellular rejection and the subsequent development of allograft vasculopathy after cardiac transplantation. J Heart Lung Transplant 2009;28:320-7.

7. Brunner-La Rocca HP, Schneider J, Künzli A, Turina M, Kiowski W. Cardiac allograft rejection late after transplantation is a risk factor for graft coronary artery disease. Transplantation 1998;65:538-43.

8. Baraldi-Junkins C, Levin HR, Kasper EK, Rayburn BK, Herskowitz A, Baughman KL. Complications of endomyocardial biopsy in heart transplant patients. J Heart Lung Transplant 1993;12:63-7.

9. Williams MJ, Lee MY, DiSalvo TG, et al. Biopsy-induced flail tricuspid leaflet and tricuspid regurgitation following orthotopic cardiac transplantation. Am J Cardiol 1996;77:1339-44.

10. Oldham N, Ott RA, Allen BA, Fopiano P, Dwyer M. Ventricular fibrillation complicating endomyocardial biopsy of a cardiac allograft. Cathet Cardiovasc Diagn 1991;23:300-1.

11. Bhat G, Burwig S, Walsh R. Morbidity of endomyocardial biopsy in cardiac transplant recipients. Am Heart J 1993; 125:1180-1.

12. Hamour IM, Burke MM, Bell AD, Panicker MG, Banerjee R, Banner NR. Limited utility of endomyocardial biopsy in the first year after heart transplantation. Transplantation 2008;85:969-74.

13. Horwitz PA, Tsai EJ, Putt ME, et al. Detection of cardiac allograft rejection and response to immunosuppressive therapy with peripheral blood gene expression. Circulation 2004;110:3815-21.

14. Deng MC, Eisen HJ, Mehra MR, et al. Noninvasive discrimination of rejection in cardiac allograft recipients using gene expression profiling. Am J Transplant 2006;6:150-60.

15. Starling RC, Pham M, Valantine H, et al. Molecular testing in the management of cardiac transplant recipients: initial clinical experience. J Heart Lung Transplant 2006;25:1389-95. [Erratum, J Heart Lung Transplant 2007;26:204.]

16. Pham MX, Deng MC, Kfoury AG, Teuteberg JJ, Starling RC, Valantine H. Molecular testing for long-term rejection surveillance in heart transplant recipients: design of the Invasive Monitoring Attenuation Through Gene Expression (IMAGE) trial. J Heart Lung Transplant 2007;26:808-14.

17. Taylor DO, Edwards LB, Boucek MM, et al. Registry of the International Society for Heart and Lung Transplantation: twenty-second Official Adult Heart Transplant Report — 2005. J Heart Lung Transplant 2005;24:945-55.

18. Mills RM, Naftel DC, Kirklin JK, et al. Heart transplant rejection with hemodynamic compromise: a multiinstitutional study of the role of endomyocardial cellular infiltrate: Cardiac Transplant Research Database. J Heart Lung Transplant 1997;16:813-21.

19. Marboe CC, Billingham M, Eisen H, et al. Nodular endocardial infiltrates (Quilty lesions) cause significant variability in diagnosis of ISHLT Grade 2 and 3A rejection in cardiac allograft recipients. J Heart Lung Transplant 2005;24:Suppl: S219-S226.

20. Winters GL, Loh E, Schoen FJ. Natural history of focal moderate cardiac allograft rejection: is treatment warranted? Circulation 1995;91:1975-80.

21. White JA, Guiraudon C, Pflugfelder PW, Kostuk WJ. Routine surveillance myocardial biopsies are unnecessary beyond one year after heart transplantation. J Heart Lung Transplant 1995;14:1052-6.

22. Taylor DO, Stehlik J, Edwards LB, et al. Registry of the International Society for Heart and Lung Transplantation: twenty-sixth Official Adult Heart Transplant Report — 2009. J Heart Lung Transplant 2009;28:1007-22.

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Pham MX, Teuteberg JJ, Kfoury AG, et al. Gene-expression profiling for rejection surveillance after cardiac transplantation. N Engl J Med 2010;362:1890-900. DOI: 10.1056/NEJMoa0912965.

Gene Expression Profiling for Rejection Surveillance After Cardiac Transplantation

SUPPLEMENTARY APPENDIX

Table of Contents

SECTION 1: SUPPLEMENTARY METHODS

1.1 Patients: Non-consecutive heart transplant recipients from 13 U.S. centers were assessed for eligibility between January 2005 and October 2009. The full list of inclusion and exclusion criteria is presented in Supplementary Table 1. Additionally, some patients who met eligibility criteria were not enrolled due to the preference of the investigator or treating physician for biopsy-based rejection monitoring. A subset of these patients may have been considered to be at higher risk for rejection on the basis of their medical history or clinical assessment. The institutional review board at each participating center approved the study.

1.2 Biopsy: Patients in the biopsy group underwent protocol surveillance biopsies at pre-specified and center-specific intervals (refer to Supplementary Table 2). Patients in the gene-expression profiling group underwent biopsies if the gene-expression profiling score was above threshold (see *Thresholds used in the IMAGE study* below). However, patients with consistently elevated gene-expression profiling scores on three consecutive study visits who did not have clinical manifestations of graft dysfunction and who did not have treatable rejection on two consecutive biopsies over a period of 3 to 9 months could be managed without a biopsy on the third or subsequent visit if there was no statistically significant increase in the score during those visits. Two scores were considered statistically different from one another if the previous score did not fall within the 95% confidence interval of the current score.

Patients in both groups underwent endomyocardial biopsy if signs or symptoms of rejection or allograft dysfunction were present at the time of the clinic visit, or if the echocardiogram showed a proportional left ventricular ejection fraction decrease of $\geq$25% compared to the first visit (reference) value. Clinically driven biopsies were permitted in both groups if signs or symptoms of heart failure developed between routine surveillance visits. Following a treated rejection episode, all patients underwent surveillance endomyocardial biopsies per center protocol, regardless of randomization arm, for a period of 2-3 event-free

months. Biopsy specimens were interpreted by the transplant center pathologist in accordance with ISHLT criteria and without knowledge of the patient's gene-expression profiling score.[1]

1.3 Gene-expression profiling: Gene-expression profiling is based on the analysis of peripheral blood mononuclear cell messenger ribonucleic acid (mRNA) using real-time polymerase chain reaction (PCR) technology. Previous studies have shown that certain gene-expression profiles correlate with histologic rejection on the endomyocardial biopsy, and their analysis may provide a valuable tool for the non-invasive monitoring of acute rejection.[2, 3, 4]

In the Cardiac Allograft Rejection Gene Expression Observation (CARGO) study, 9 centers enrolled 737 patients and collected clinical data, peripheral blood mononuclear cell samples, and biopsy specimens during 5834 routine and non-routine clinical encounters.[4] The primary objective of the CARGO study was to develop and validate a gene-expression profile test for acute cellular rejection. A linear discriminant equation (classifier) was derived by sequentially fitting the gene-expression data from 145 peripheral blood samples to maximize the agreement with the histology classification on the corresponding biopsy samples. The final classifier, yielding a score between 0 and 40, combined the expression levels of 11 informative genes which best distinguished rejection (ISHLT biopsy grade ≥3A) from non-rejection (ISHLT biopsy grade <3A). An additional 9 genes are included in the test for control and normalization purposes.

The gene-expression profiling test is available commercially as an FDA cleared in vitro diagnostic multivariate index assay (AlloMap®; XDx, Brisbane, CA). The test is intended to aid in the identification of heart transplant recipients with stable allograft function who have a low probability of moderate or severe acute cellular rejection at the time of testing, in conjunction with standard clinical assessment. The AlloMap® test is approved for heart transplant recipients 15 years of age or older who are at least two months (≥55 days) post transplantation.

Genes and pathways: The 11 informative genes in the AlloMap® test are involved in pathways hypothesized to play a role in immune activation during acute cellular rejection, including T cell priming, platelet activation, proliferation and mobilization of immature erythrocytes, and steroid responsiveness.[4] Supplementary Table 3 lists the individual genes, their patterns of expression during rejection, predominant source of expression in blood, and their postulated role in immune activation and rejection.

Performance characteristics: Supplementary Table 4 presents the performance characteristics for the AlloMap® test across a range of score thresholds in the 2 to 6 month and >6 month post-transplantation periods. The test characteristics reported here differ slightly from previously published reports and reflect updated data utilized for FDA clinical validation studies in 2008. Test characteristics were derived for an independent 300 samples from 154 patients in the CARGO study that were not used in the discovery and development of the classifier.

Specimen processing and reporting of test results: The AlloMap® blood samples were processed locally and shipped frozen to the XDx laboratory. An AlloMap® score from 0-40 was reported to the transplant center within 4 days of specimen collection.

Thresholds used in the IMAGE study: The initial protocol specified a gene-expression profiling score of 30 or higher to prompt a required endomyocardial biopsy. This threshold was selected based upon the initial findings from the CARGO study, showing that a score below 30 was associated with a negative predictive value of 99.6% for concurrent ISHLT Grade 3A (2R) or higher rejection. On November 7, 2005, a protocol amendment increased the threshold to 34 to minimize the number of biopsies needed in the gene-expression profiling group without compromising the assay performance. The decision to increase the threshold was based upon additional analyses from the CARGO study showing that the negative predictive value of the gene-expression profiling test remained robust (99.2%) at a higher threshold of 34 while reducing the number of positive tests from 50.8% to 22.3%.[5] The IMAGE investigators

recognized that the higher threshold represented a more pragmatic yet still prudent threshold in order to maintain a low risk of missing asymptomatic rejection episodes while reducing the number of unnecessary biopsies.

1.4 Rejection therapy: Treated rejection was defined as the administration of anti-rejection therapy such as pulse steroids, antibody therapy, or plasmapheresis, with or without histological findings of rejection on the endomyocardial biopsy. Rejection therapy was given based upon endomyocardial biopsy results and other conventional diagnostic testing according to center specific practices, which are summarized in Supplementary Table 9.

1.5 Primary end point:

Definition: The primary outcome was a composite of the following subcomponents: rejection with hemodynamic compromise, graft dysfunction due to other causes, death, or retransplantation. The definitions are presented below:

1. **Rejection with hemodynamic compromise:** Criteria (a) **and** either criteria (b) or (c) must be met.
    a. Presence of hemodynamic compromise: The presence of one or more of the following criteria is required:
        i. Absolute drop in LVEF ≤30% at the time of the rejection episode, as confirmed by the Core Echocardiography Laboratory.
        ii. Proportional decrease in LVEF ≥25% compared to the reference (first study visit) value at the time of the rejection episode, as confirmed by the Core Echocardiography Laboratory.
        iii. Cardiac index < 2 L/min/m^2 at the time of the rejection episode.
        iv. Use of inotropic drugs to support circulation at any time during the rejection episode. Use of dopamine at ≤3 mcg/kg/min, when used to enhance renal perfusion, did not count toward this criteria.

b. Supporting histologic or immunologic evidence of rejection, as determined by the local pathologist's review. At least one of the criteria below is required:

   i. Cellular rejection of ISHLT Grade 3A (1990 classification) or Grade 2R (2004 classification).

   ii. Antibody-mediated rejection of ISHLT Grade AMR1 (2004 classification).

   iii. Antibody-mediated rejection as defined by histologic evidence of capillary injury and/or positive immunopathologic evidence of antibody mediated injury (immunofluorescence or immunohistochemistry).

   iv. Mixed cellular and antibody-mediated rejection.

c. Probable rejection: This category was used by the Endpoints Committee to classify events considered to be secondary to rejection in the absence of histologic confirmation on biopsy. An example would include a patient who presents with hemodynamic compromise, no evidence of cardiac allograft vasculopathy or histologic evidence of rejection, and whose condition improves with initiation of anti-rejection therapy.

2. **Graft dysfunction due to other causes.** Criteria (a) **and** either criteria (b), (c), or (d) must be met.

   a. Presence of hemodynamic compromise: The definition is the same as listed above.

   b. Cardiac allograft vasculopathy: The diagnosis requires any one of the following criteria, either prior to or at the time of the event:

      i. Stenosis of ≥50% within any major epicardial coronary vessel or branches on angiography.

      ii. Severe diffuse or distal vessel tapering on angiography.

      iii. Maximal intimal thickness ≥ 0.5 mm in any major epicardial coronary vessel at the time of intravascular ultrasound.

iv. Evidence of significant intimal hyperplasia at the level of the arterioles or intra-myocardial small vessels at autopsy.

v. Evidence of recent ischemic injury on biopsy.

c. Probable cardiac allograft vasculopathy: This category was used by the Endpoints Committee to classify events that are related to cardiac allograft vasculopathy but that do not meet the diagnostic criteria above. An example would include a patient with graft dysfunction and epicardial stenosis of <50% who does not respond to empiric anti-rejection therapy and for which no autopsy information is available.

d. Nonspecific or other: This category was coded if the etiology of graft dysfunction could not be readily classified into one of the two previous categories.

3. **Death or retransplantation.** Death or retransplantation from any cause was considered.

Rationale for selection of primary end point: We selected a composite primary end point to include events whose detection, if present, would not be influenced by the specific rejection monitoring method used. For example, rejection with hemodynamic compromise would manifest clinically with symptoms or echocardiographic evidence of graft dysfunction in both the gene-expression profiling and biopsy groups. Since histologic sampling for rejection was performed, by study design, less frequently in the gene-expression profiling group, our end point would also need to identify the sequelae of undetected rejection episodes. Since rejection in the late (>1 year) post-transplantation can theoretically cause graft dysfunction through progressive myocardial fibrosis of intimal hyperplasia of the coronary vasculature, we included graft dysfunction not associated with rejection. Finally, since graft dysfunction can rapidly progress to death or the need for retransplantation, these events were also included in the primary end point.

1.6 Quality of life and patient satisfaction: We assessed the effects of rejection monitoring strategy on patients' quality of life using the Medical Outcomes Study 12-item Short Form (SF-12) General Health Survey (version 2). The SF-12 survey measures health status in eight core domains, and the results are expressed in terms of two meta-scores. The Physical Component Summary assesses physical functioning, bodily pain, physical role functioning, and general health, while the Mental Component Summary assesses vitality, social functioning, emotional role functioning, and mental health. The SF-12 summary scores for mental health and physical health have a range of 0 to 100, with higher scores indicating better functioning. They were designed to have a mean score of 50 and a standard deviation of 10 in a representative sample of the U.S. population. We also assessed patient satisfaction with the method of rejection monitoring using an ordinal scale ranging from 1 (very unhappy) to 10 (very happy). The SF-12 health survey and patient satisfaction questionnaire were administered at enrollment, 1 year in the study, and at the completion of 2 years in the study.

Supplementary Table 1: IMAGE eligibility criteria

Inclusion criteria
1. Heart transplant recipients who are between >6 months to 5 years (>6-60 months) post-transplantation. 2. Age ≥18 years. 3. Stable outpatient being seen for routine monitoring of rejection. Stability is defined as absence of prior or current evidence of either severe cardiac allograft vasculopathy (CAV) or antibody-mediated rejection (AMR) with associated hemodynamic compromise. a. Severe CAV is defined as either A) >50% left main stenosis; B) ≥50% stenosis in ≥2 primary vessels (proximal 1/3 or middle 1/3 of the LAD or LCx, RCA to takeoff of PDA in right-dominant coronary circulations) or C) isolated branch stenoses of >50% in all 3 systems (diagonal branches, obtuse marginal branches, distal 1/3 of LAD or LCx, PDA, PLB, and RCA to takeoff of PDA in non-dominant systems). b. AMR with associated hemodynamic compromise is defined as AMR (defined according to local criteria) with either A) a left ventricular ejection fraction (LVEF) ≤ 30% or at least 25% lower than the baseline value, B) a cardiac index < 2 L/min/m^2, or C) the use of inotropic agents to support circulation. 4. Left ventricular ejection fraction ≥45% by Echocardiography, Multiple Gated Acquisition (MUGA) scan, or ventriculography at study entry (baseline / enrollment study)

Supplementary Table 1 (Continued): IMAGE eligibility criteria

Exclusion criteria
1. Patients <7 months after heart transplantation.
2. Any symptoms or clinical signs of impaired allograft function:
a. Symptoms of Congestive Heart Failure (CHF) at the enrollment visit.
b. Signs of decompensated heart failure, including the development of a new S3 gallop at the enrollment visit.
c. Elevated right heart pressures with diminished cardiac index <2.2 L/min/m^2 that is new compared to a previous measurement within 6 months.
d. Decrease in LVEF as measured by echocardiography: ≥25% compared to prior measurement within 6 months.
3. Rejection therapy for biopsy-proven ISHLT Grade 3A or higher during the preceding 2 months.
4. Major changes in immunosuppression therapy within previous 30 days (e.g., discontinuation of calcineurin inhibitors, switch from mycophenolate mofetil to sirolimus or vice versa).
5. Unable to give written informed consent.
6. Patient receiving hematopoietic growth factors (e.g. Neupogen, Epogen) currently or during the previous 30 days.
7. Patients receiving ≥ 20 mg/day of prednisone equivalent corticosteroids at the time of enrollment
8. Patient enrolled in a trial requiring routine surveillance endomyocardial biopsies.
9. Patient received transfusion within preceding 4 weeks.
10. Patients with end-stage renal disease requiring some form of renal replacement therapy (hemodialysis or peritoneal dialysis).
11. Pregnancy at the time of enrollment.

Supplementary Table 2: Rejection surveillance schedule at IMAGE centers

Study Center	>6-12 months	Year 2	Year 3	Year 4	Year 5
Intermountain Medical Center, UT	7, 8.5, 10, 12	3, 6, 9, 12	4, 8, 12	6, 12	12
University of Chicago Medical Center, IL	7, 8, 9, 10, 11, 12	3, 6, 12	12	12	12
Hospital of the University of Pennsylvania, PA	7, 8, 9, 10, 11, 12	3, 6, 9, 12	6, 12	6, 12	6, 12
St Luke's Hospital, Kansas, MO*	7, 8, 9, 10, 11, 12	3, 6, 9, 12	6, 12	—	—
Barnes Jewish Hospital, MO	7, 8, 9, 10, 11, 12	3, 6, 9, 12	6, 12	12	12
Columbia University Medical Center, NY	8, 10, 12	3, 6, 9, 12 or 4, 8, 12	4, 8, 12 or 12	6, 12 or 12	6, 12 or 12
Cleveland Clinic, OH	8, 10, 12	3, 6, 9, 12	3, 6, 9, 12	6, 12	6, 12
University of Pittsburgh Medical Center, PA*	8, 10, 12	3, 6, 9, 12	3, 6, 9, 12	12	12
VA Palo Alto Medical Center, CA†	8, 10, 12	4, 8, 12	4, 8, 12	6, 12	6, 12
Northwestern University Medical Center, IL	8, 10, 12	3, 6, 9, 12 or 4, 8, 12	6, 12	6, 12	6, 12
Stanford University Medical Center, CA†	9, 12	4, 8, 12	4, 8, 12	6, 12	6, 12

Texas Heart Institute, TX‡	9, 12	6, 12	12	12	12
Newark Beth Israel Medical Center, NJ	§	6, 12	6, 12	12	12

Footnotes for Supplementary Table 2:

* Patients >3 years post-transplant were followed-up for surveillance monitoring using echocardiogram and physical exams, instead of by biopsy. Therefore no gene-expression profiling testing was done for patients >3 years post-transplant.

† Patients >5 years post-transplant were followed-up for surveillance monitoring using echocardiogram and physical exams instead of by biopsy. Therefore no gene-expression profiling testing was done for patients >5 years post-transplant.

‡ The site did not biopsy patients who were 1 year post-transplant, therefore no gene-expression profiling testing was done after 1 year post-transplant.

§ Patients in this time frame excluded from IMAGE due to competing clinical study at this center.

Supplementary Table 3: Genes and pathways contained in AlloMap® test

AlloMap Genes	Expression Change with ACR	Predominant Source of Expression in Blood	Role in Immune Activation and Rejection
IL1R2, FLT3, ITGAM	Decreased	Monocytes	Steroid response
MARCH8, WDR40A	Increased	Reticulocytes	Proliferation and mobilization of erythrocytes
PF4, C6orf25	Decreased	Platelets	Platelet activation
RHOU	Increased	T cells and Monocytes	Unknown
PDCD1	Increased	T cells	T cell priming
ITGA4	Increased	T cells	T cell priming
SEMA7A	Increased	T cells, B cells and Immature Neutrophils	Unknown

Supplementary Table 4: AlloMap® performance metrics estimated from the CARGO patient population. Performance is given by time post-transplantation.

AlloMap Score	>2 – 6 Months (N=166 samples)					>6 Months (N=134 samples)				
	% Pts Below	PPV ≥3A(2R)	PPV Std. Err.	NPV <3A(2R)	NPV Std. Err.	% Pts Below	PPV ≥3A(2R)	PPV Std. Err.	NPV <3A(2R)	NPV Std. Err.
19	<22.4	≤2.7%	≤0.1%	100.0%	0.0%	≤5.4	≤1.8%	0.0%	100.0%	0.0%
20	24.3%	2.8%	0.2%	100.0%	0.0%	8.1%	1.8%	0.1%	100.0%	0.0%
21	33.6%	2.5%	0.4%	98.8%	0.8%	9.8%	1.9%	0.1%	100.0%	0.0%
22	38.8%	2.7%	0.5%	98.9%	0.7%	11.0%	1.9%	0.1%	100.0%	0.0%
23	41.8%	2.9%	0.5%	99.0%	0.6%	14.1%	2.0%	0.1%	100.0%	0.0%
24	47.5%	3.2%	0.6%	99.1%	0.6%	18.4%	2.1%	0.1%	100.0%	0.0%
25	56.0%	3.8%	0.7%	99.3%	0.5%	22.1%	2.2%	0.1%	100.0%	0.0%
26	61.4%	3.8%	0.9%	99.0%	0.5%	26.8%	2.3%	0.1%	100.0%	0.0%
27	63.6%	3.4%	1.0%	98.7%	0.5%	31.6%	1.9%	0.4%	98.7%	0.9%
28	68.3%	3.3%	1.1%	98.5%	0.5%	39.1%	2.1%	0.5%	98.9%	0.7%
29	73.7%	4.0%	1.3%	98.6%	0.4%	40.8%	2.1%	0.5%	99.0%	0.7%
30	77.2%	4.6%	1.6%	98.6%	0.4%	50.6%	2.1%	0.6%	98.7%	0.6%
31	81.0%	3.3%	1.6%	98.2%	0.4%	54.1%	2.3%	0.7%	98.8%	0.6%
32	85.6%	2.9%	2.0%	98.0%	0.3%	63.1%	2.9%	0.9%	99.0%	0.5%
33	89.4%	4.0%	2.7%	98.1%	0.3%	72.4%	3.8%	1.3%	99.1%	0.4%
34	91.7%	5.0%	3.5%	98.2%	0.3%	79.1%	4.1%	1.7%	98.9%	0.4%
35	94.5%	5.7%	4.8%	98.1%	0.2%	84.1%	4.0%	2.2%	98.7%	0.4%
36	97.3%	7.6%	13.8%	98.1%	0.2%	90.2%	5.4%	3.2%	98.7%	0.3%
37	97.8%	9.5%	21.1%	98.1%	0.2%	91.7%	—	–	98.4%	0.2%
38	100.0%	—	–	97.9%	0.0%	96.5%	—	–	98.2%	0.0%
39	100.0%	—	–	97.9%	0.0%	97.7%	—	–	98.3%	0.0%

Supplementary Table 5: Rejection therapy protocols at IMAGE centers

ISHLT biopsy grade	Hemodynamic compromise absent	Hemodynamic compromise present
0, 1A (1R), 1B (1R), 2 (1R)	No treatment except at Columbia University, where Grade 1B (1R) rejection during the first year-transplantation was treated with oral corticosteroids.	High dose (500-1000mg) IV corticosteroids **with or without** Cytolytic antibody therapy (OKT3, Thymoglobulin, ATGAM) **Consider empiric treatment for antibody mediated rejection**
3A (2R)	1 mg/kg oral corticosteroids with rapid taper **or** High dose (500-1000mg) IV corticosteroids x 3 days	High dose (500-1000mg) IV corticosteroids x 3 days **with or without** Cytolytic antibody therapy
3B, 4 (3R)	High dose (500-1000mg) IV corticosteroids **with or without** Cytolytic antibody therapy (Thymoglobulin 1.5 mg/kg up to 120 mg daily x 3 days)	High dose (500-1000mg) IV corticosteroids **with or without** Cytolytic antibody therapy
AMR positive	No treatment	Apheresis **with or without** IV immune globulin High dose IV (500-1000mg) corticosteroids Rituximab infusion

SECTION 2: SUPPLEMENTARY RESULTS

2.1 Patient population: The baseline characteristics of the IMAGE cohort is compared against a cohort of pediatric and adult heart transplant recipients in 2007 as reported to the Organ Procurement and Transplantation Network (OPTN) and United Network of Organ Sharing (UNOS) (See Supplementary Table 6). Compared to the OPTN/UNOS cohort, the IMAGE cohort contained a greater proportion of men and a lower proportion of African-American patients. The higher proportion of patients with coronary artery disease and lower proportion of patients with congenital heart disease in the IMAGE population likely reflects the exclusion of pediatric patients. The use of induction therapy was also higher in the IMAGE cohort, particularly with respect to the use of antithymocyte globulin and interleukin-2 receptor antagonists.

2.2 Immunosuppression: Overall immunosuppression intensity was similar in the gene-expression profiling and biopsy arms throughout the study. The mean cyclosporine 12-hour trough level was higher in the gene-expression profiling group compared to the biopsy group at baseline (177 ng/mL vs. 141 ng/mL, P=0.02); however, the difference between the groups narrowed during the study such that the mean drug levels were similar when averaged throughout the study (142 ng/mL in the gene-expression profiling group vs. 131 ng/mL in the biopsy group, P=0.28). The mean tacrolimus drug levels in the gene-expression profiling group, compared to the biopsy group, was numerically higher at study entry (8.8 ng/mL vs. 8.1 ng/mL, P=0.08) and throughout the study (8.1 ng/mL vs. 7.6 ng/mL, P=0.06), but the differences were marginally statistically significant and not clinically meaningful.

2.3 Gene-expression profiling scores and biopsy results: The mean test score for patients in the gene-expression profiling arm was 29.9 ± 4.9. Among the 1190 gene-expression profiling scores reported, 302 (25%) of scores were ≥34 (see Supplementary Figure 1). Biopsies were

performed in conjunction with 274 (91%) elevated scores, either in response to the elevated score in asymptomatic patients (265 biopsies) or due to the presence of both an elevated score and clinical evidence of graft dysfunction (9 biopsies). A biopsy was not required by the protocol in 28 instances (9%) of persistently elevated gene-expression profiling scores occuring in patients with no history of rejection on prior biopsies.

Among the 265 biopsies performed in response to elevated gene-expression profiling scores (≥34), 143 (54%) biopsies revealed no evidence of rejection (ISHLT Grade 0), 111 (42%) biopsies revealed ISHLT Grade 1R rejection, 8 (3%) biopsies revealed ISHLT Grade 2R rejection, and 3 (1%) biopsy revealed ISHLT Grade 3R rejection.

2.4 Rejection rates: There were 81 discrete treated rejection episodes (34 in the gene-expression profiling group and 47 in the endomyocardial biopsy group) observed in 61 patients. Among patients in the gene-expression profiling group, 20 (59%) treated rejection episodes were prompted by overt heart failure and/or by echocardiographic evidence of graft dysfunction, 7 (21%) episodes were associated with both clinical manifestations and elevated gene-expression profiling scores, and 6 asymptomatic episodes (18%) were detected solely on the basis of elevated gene-expression profiling scores. In contrast, 22 (47%) of the treated rejection episodes in the endomyocardial biopsy group were asymptomatic and detected by routine endomyocardial biopsy (see Supplementary Table 9).

The incidence of acute cellular rejection (ISHLT grades 2R or 3R) or mixed rejection (acute cellular rejection with antibody-mediated rejection) in our study was 6.1% of biopsies (5.6% of patients) in the gene-expression profiling group and 3.0% of biopsies (9.3% of patients) in the endomyocardial biopsy group. Antibody-mediated rejection was observed in 6.1% of biopsies (5.2% of patients) in the gene-expression profiling group and in 1.7% of biopsies (5.1% of patients) in the endomyocardial biopsy group.

Supplementary Table 6: Baseline characteristics of the IMAGE study population compared to OPTN/UNOS pediatric and adult heart transplant recipients from January 1, 2007 – December 31, 2007*

	IMAGE Study Cohort	UNOS/OPTN Cohort
	(N=602)	**(N=2207)**
Age at transplant – %		
12 – 17 years	0.8	4.8
18 – 34 years	11.1	10.6
35 – 49 years	20.8	19.9
50 – 64 years	52.5	43.5
65+ years	14.8	11.1
Male gender – (%)	81.9	73.7
Race – (%)		
White	77.7	67.5
Hispanic	6.5	9.4
African American	11.8	19.3
Asian or Pacific Islander	2.2	3.2
Other	1.8	0.7
Indication for heart transplantation – no. (%)		
Coronary artery disease	42.7	35.6
Non-ischemic cardiomyopathy	51.0	53.4
Valvular heart disease	1.8	1.5
Congenital heart disease	3.0	7.9
Graft vasculopathy / Retransplant	0.7	Not available

Other	0.8	1.4
Induction therapy use – (%)		
OKT3	1.5	3.2
Antithymocyte globulin	17.4	4.2
Basiliximab	12.1	14.2
Daclizumab	21.4	12.7
Alemtuzumab	4.2	2.6
Other	1.3	--
Immunosuppression – (%)†		
Cyclosporine	26.9	33.0
Tacrolimus	72.4	64.9
Mycophenolate mofetil / mycophenolic acid	80.9	85.6
Azathioprine	6.8	2.8
Sirolimus	19.6	1.2
Prednisone	42.2	88.7

Footnotes for Supplementary Table 6:

* Data obtained from www.ustransplant.org

† Denotes all medications while in IMAGE study. Note: Medications in the UNOS/OPTN cohort reflect maintenance immunosuppression use prior to discharge from the transplant hospitalization.

Supplementary Table 7: Selected adverse events during the study

Event	Gene profiling	Biopsy	P value
	(N=297)	(N=305)	
Cardiac*	7 (2.4)	8 (2.6)	0.12
Angina pectoris	0	1 (0.3)	
Arrhythmia, palpitations, or tachycardia	3 (1.0)	3 (1.0)	
Congestive heart failure	1 (0.3)	1 (0.3)	
Cardiomegaly	0	1 (0.3)	
Coronary artery disease	2 (0.7)	1 (0.3)	
Pericardial effusion	1 (0.3)	1 (0.3)	
Infections	53 (17.8)	43 (14.1)	0.22
Bacterial	6 (2.0)	3 (1.0)	
Fungal	0	1 (0.3)	
Viral (cytomegalovirus, herpes simplex, herpes zoster)	9 (3.0)	3 (1.0)	
Viral (other)	5 (1.7)	6 (2.0)	
Other / unspecified	33 (11.1)	30 (9.8)	
Neoplasms	11 (3.7)	10 (3.3)	0.83
Skin cancer	4 (1.3)	4 (1.3)	
Squamous cell carcinoma	1 (0.3)	2 (0.7)	
Basal cell carcinoma	1 (0.3)	0	
Malignant melanoma	1 (0.3)	2 (0.7)	
Unspecified	1 (0.3)	0	
Breast cancer	1 (0.3)	0	

Colon cancer	1 (0.3)	1 (0.3)
Lung cancer	1 (0.3)	1 (0.3)
Prostate cancer	4 (1.3)	3 (1.0)
Pituitary adenoma	0	1 (0.3)

Footnote for Supplementary Table 7:

* Not meeting endpoint definition and not biopsy-related.

Supplementary Table 8: Biopsy use during study*

	Gene-expression Profiling	Biopsy
	(N=287)	(N=292)
Total biopsies -- no. biopsies	409	1249
Routine per-protocol surveillance	N/A	1125
Performed due to elevated GEP[†]	274	N/A
Clinically driven[‡]	70	31
Performed within 90 days of rejection treatment	52	58
Off-protocol[§]	13	35
Frequency of biopsies per patient year of follow-up – no. patients (%)		
0 biopsies/patient year[‖]	142 (50)	4 (1.4)
1 – 2 biopsies/patient year	108 (38)	94 (32)
3 – 4 biopsies/patient year	21 (7.3)	135 (46)
≥ 5 biopsies/patient year	16 (5.6)	59 (20)
Number of biopsies per patient year of follow-up stratified by time post-transplantation – median (min – max)		
Overall	0.5 (0.0 – 15.9)	3.0 (0.0 – 22.1)
6 – 12 months post-transplantation	0.7 (0.0 –15.9)	5.1 (2.8 – 22.1)
12 – 36 months post-transplantation	0.5 (0.0 – 12.0)	3.0 (0.0 – 7.5)
36 – 60 months post-transplantation	0.2 (0.0 – 7.1)	1.9 (0.0 – 4.5)

Footnotes for Supplementary Table 8:

* The analyses include patients who completed at least 1 study visit and were followed for a minimum of 30 days in the study. Both scheduled study visits and non-scheduled outpatients visits were included.

† As mandated by the study protocol. 9 biopsies were performed for scores ≥30 prior to the protocol amendment on November 7, 2005. 265 biopsies were performed for scores ≥34.

‡ Clinically driven biopsies were performed, per-protocol, for clinical signs or symptoms of congestive heart failure or for graft dysfunction, defined by a decrease in the LVEF of ≥25% compared to the first visit (reference) value.

§ Off protocol biopsies included additional biopsies performed that were not mandated by the protocol and biopsies for which no reason was given.

‖ The number of patients who had exactly 0 biopsies on-study was 133 in the gene-expression profiling group and 2 in the biopsy group. An additional 9 patients in the gene-expression profiling and 2 patients in the biopsy group were also included in this category because they had 0.49 biopsies per patient year of follow-up, which was rounded down to 0 per year.

Supplementary Table 9: Treated rejection episodes

	Gene-expression Profiling	Biopsy
	(N=34 Episodes)	(N=47 Episodes)
Biopsy Histology		
Grade 0	2	5
Grade 1R rejection	8	8
Grade 2R rejection	11	22
Grade 3R rejection	2	3
Antibody-mediated rejection	8	6
Mixed rejection	3	3
Presentation		
Clinical signs or symptoms	6	8
Graft dysfunction on echocardiogram	2	7
Clinical signs/symptoms and graft dysfunction on echocardiogram	12	10
Clinical signs/symptoms and elevated GEP score	6	N/A
Clinical signs/symptoms, graft dysfunction on echocardiogram, and elevated GEP score	1	N/A
Elevated GEP score	6	N/A
Asymptomatic	0	22
Other*	1	0

Footnote for Supplementary Table 9:

* One patient underwent endomyocardial biopsy due to a rising gene-expression profiling score that did not meet the threshold for biopsy (protocol violation).

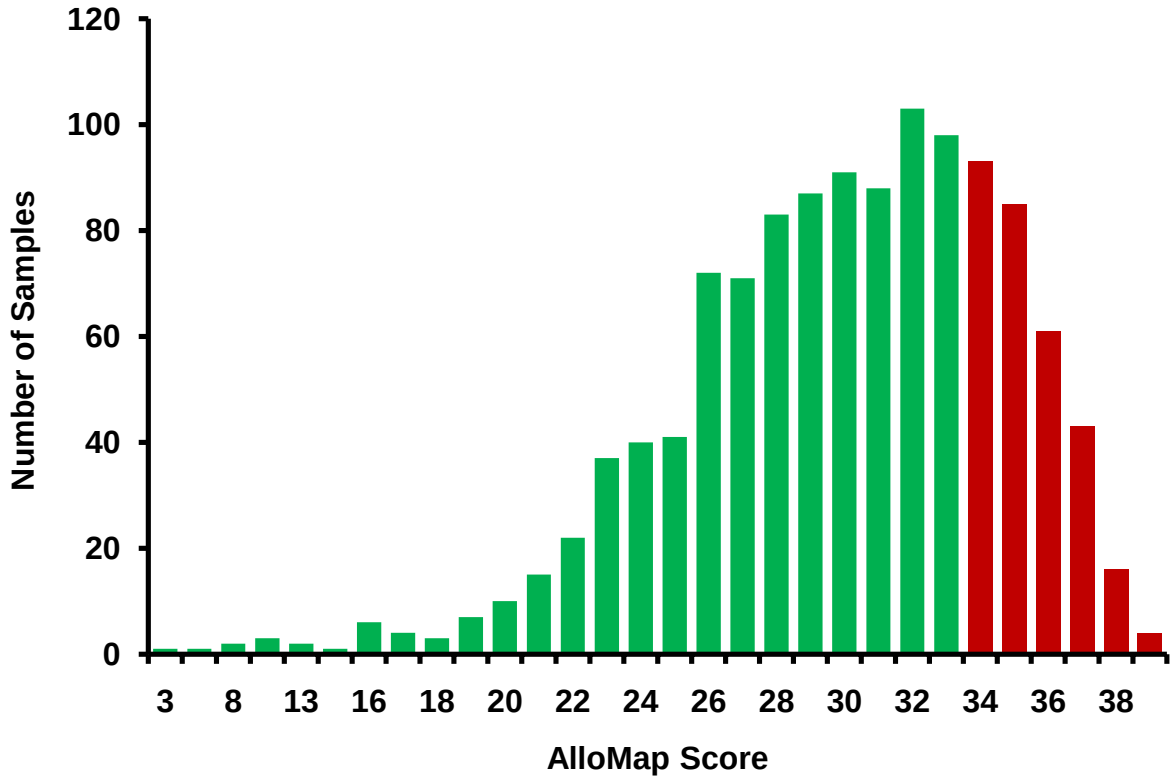
Supplementary Table 10: Quality of life and patient satisfaction scores.

	Gene-expression Profiling			Biopsy			P-values*		
	Enrollment	Year 1	Year 2	Enrollment	Year 1	Year 2	Enrollment	Year 1	Year 2
Patients on-study	N = 297	N = 209	N = 101	N = 305	N = 211	N = 91			
Patients completing SF-12 Survey	N = 249	N = 148	N = 89	N = 239	N = 146	N = 83			
SF-12 Mental health summary score	51.6±10.1	50.3±10.8	50.8±10.1	52.4±8.9	51.7±9.7	50.7±9.8	0.33	0.23	0.66
SF-12 Physical health summary score	45.5±10.6	44.7±11.4	45.1±11.6	46.8±9.0	47.3±9.6	46.2±10.9	0.13	0.03	0.52
Patients completing satisfaction questionnaire	N = 269	N = 153	N = 92	N = 263	N = 155	N = 91			
Patient satisfaction score	6.86±2.75	8.15±2.95	8.74±1.90	6.74±2.71	6.64±2.98	6.66±2.81	0.61	<0.001	<0.001

Footnote for Supplementary Table 10:

* P values were obtained from the two-sample t-test.

Supplementary Figure 1: Distribution of AlloMap® scores in the gene-expression profiling group



SECTION 3: REFERENCES

1. Stewart S, Winters GL, Fishbein MC, et al. Revision of the 1990 working formulation for the standardization of nomenclature in the diagnosis of heart rejection. *J Heart Lung Transplant.* Nov 2005;24(11):1710-1720.

2. Horwitz PA, Tsai EJ, Putt ME, et al. Detection of cardiac allograft rejection and response to immunosuppressive therapy with peripheral blood gene expression. *Circulation.* Dec 21 2004;110(25):3815-3821.

3. Schoels M, Dengler TJ, Richter R, Meuer SC, Giese T. Detection of cardiac allograft rejection by real-time PCR analysis of circulating mononuclear cells. *Clin Transplant.* Oct 2004;18(5):513-517.

4. Deng MC, Eisen HJ, Mehra MR, et al. Noninvasive discrimination of rejection in cardiac allograft recipients using gene expression profiling. *Am J Transplant.* Jan 2006;6(1):150-160.

5. Starling RC, Pham M, Valantine H, et al. Molecular testing in the management of cardiac transplant recipients: initial clinical experience. *J Heart Lung Transplant.* Dec 2006;25(12):1389-1395.

6. Pham MX, Deng MC, Kfoury AG, Teuteberg JJ, Starling RC, Valantine H. Molecular testing for long-term rejection surveillance in heart transplant recipients: design of the Invasive Monitoring Attenuation Through Gene Expression (IMAGE) trial. *J Heart Lung Transplant.* Aug 2007;26(8):808-814.